

## PATENT ABSTRACTS OF JAPAN

(11) Publication number: **10124291 A**(43) Date of publication of application: **15.05.98**

(51) Int. Cl.  
**G06F 3/16**  
**G10L 3/00**  
**G10L 3/00**  
**H04B 7/26**  
**H04M 3/42**

(21) Application number: **08274452**(71) Applicant: **CASIO COMPUT CO LTD**(22) Date of filing: **17.10.96**(72) Inventor: **YAMAKITA TORU**

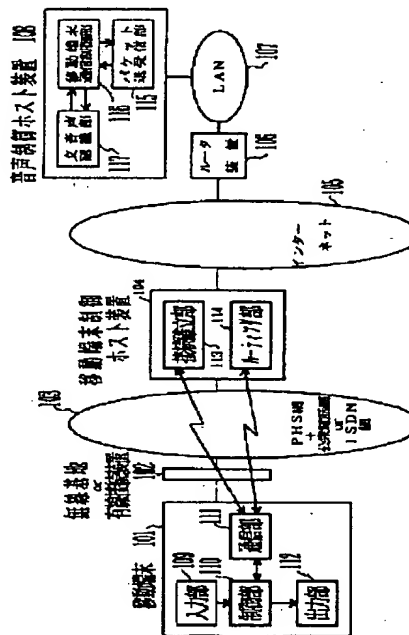
(54) **SPEECH RECOGNITION COMMUNICATION  
 SYSTEM FOR MOBILE TERMINAL**

## (57) Abstract:

**PROBLEM TO BE SOLVED:** To realize a speech recognition function as a user interface of a mobile terminal in communication environment using the mobile terminal at practical accuracy and cost.

**SOLUTION:** In the mobile terminal 101, a speech signal inputted from an input part 109 in a call or an off-line state is transmitted from a control part 110 and a communicating part 111 to a PHS network 103, from which the speech signal is transmitted to a speech control host device 108 with a mobile terminal control host device 104 and an internet 105. The speech signal is received by a mobile terminal communication control part 116 with a packet transmitting/receiving part 115 in the speech control host device 108 and recognized by a sentence/voice recognizing part 117. Recognized speech/sentence data obtained as the result is returned in block to the mobile terminal 101 in real time or a desired timing, received by the control part 110 with the communication part 111 and displayed in an output part 112.

COPYRIGHT: (C)1998,JPO



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-124291

(43) 公開日 平成10年(1998) 5月15日

(51) Int. Cl.<sup>6</sup>

識別記号

F I

G 0 6 F 3/16

3 1 0

G 0 6 F 3/16

3 1 0 A

G 1 0 L 3/00

5 5 1

G 1 0 L 3/00

5 5 1 A

5 6 1

5 6 1 H

H 0 4 B 7/26

H 0 4 M 3/42

P

H 0 4 M 3/42

H 0 4 B 7/26

Q

審査請求 未請求 請求項の数7 O L (全 29 頁)

(21) 出願番号

特願平8-274452

(22) 出願日

平成8年(1996)10月17日

(71) 出願人 000001443

カシオ計算機株式会社

東京都渋谷区本町1丁目6番2号

(72) 発明者 山北 徹

東京都羽村市栄町3丁目2番1号 カシオ

計算機株式会社羽村技術センター内

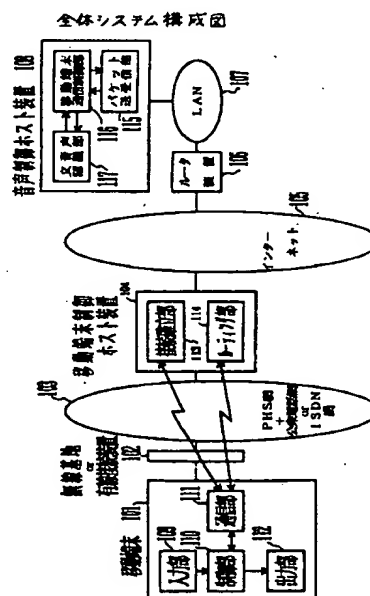
(74) 代理人 弁理士 阪本 紀康

(54) 【発明の名称】 移動端末音声認識通信システム

(57) 【要約】

【課題】 移動端末を用いた通信環境において、そのユーザインタフェースとしての音声認識機能を実用的な精度及びコストで実現することにある。

【解決手段】 移動端末101において、通話時に又はオフライン状態で入力部109から入力された音声信号が、制御部110及び通信部111からPHS網103に送出され、そこから移動端末制御ホスト装置104及びインターネット105を介して、音声制御ホスト装置108に送られる。この音声信号は、同装置内のパケット送受信部115を介して移動端末通信制御部116で受信され、文音声認識部117で認識される。その結果得られる認識音声文章データは、リアルタイムに又は所望のタイミングで一括して移動端末101に返送され、通信部111を介して制御部110で受信され、出力部112で表示される。



## 【特許請求の範囲】

【請求項1】 移動端末がホスト装置と通信する通信システムであって、前記移動端末内に、無線網又は有線網の何れか一方又は両方から構成される中継網を介して間接的に又は該中継網を介さずに直接的に前記ホスト装置である音声制御ホスト装置に接続するホスト接続手段と、

音声を入力する音声入力手段と、

前記ホスト接続手段による接続動作の後、前記音声入力手段から入力される音声データを前記音声制御ホスト装置に送信する音声データ送信手段と、

前記音声制御ホスト装置から返信される認識音声データを受信する認識音声データ受信手段と、

該受信された認識音声データを表示又は編集する認識音声データ表示／編集手段と、

を含み、

前記音声制御ホスト装置内に、

前記移動端末内のホスト接続手段による接続動作にตอบสนองして、前記移動端末を識別して接続する移動端末接続手段と、

現在接続されている移動端末毎に、前記音声データを受信する音声データ受信手段と、

現在接続されている移動端末毎に、前記音声データ受信手段により受信された音声データに対して音声認識処理を実行する音声認識手段と、

現在接続されている移動端末毎に、前記音声認識手段による音声認識処理によって得られる認識音声データを、それに対応する移動端末に返信する認識音声データ返信手段と、

を含む、

ことを特徴とする移動端末音声認識通信システム。

【請求項2】 移動端末がホスト装置と通信する通信システムに使用される前記移動端末であって、無線網又は有線網の何れか一方又は両方から構成される中継網を介して間接的に又は該中継網を介さずに直接的に前記ホスト装置である音声制御ホスト装置に接続するホスト接続手段と、

音声を入力する音声入力手段と、

前記ホスト接続手段による接続動作の後、前記音声入力手段から入力される音声データを前記音声制御ホスト装置に送信する音声データ送信手段と、

前記音声制御ホスト装置から返信される認識音声データを受信する認識音声データ受信手段と、

該受信された認識音声データを表示又は編集する認識音声データ表示／編集手段と、

を含むことを特徴とする移動端末。

【請求項3】 移動端末がホスト装置と通信する通信システムに使用される前記ホスト装置であって、無線網又は有線網の何れか一方又は両方から構成される中継網を介して間接的に又は該中継網を介さずに直接的

に前記移動端末が実行する接続動作にตอบสนองして、前記移動端末を識別して接続する移動端末接続手段と、

現在接続されている移動端末毎に、音声データを受信する音声データ受信手段と、

現在接続されている移動端末毎に、前記音声データ受信手段により受信された音声データに対して音声認識処理を実行する音声認識手段と、

現在接続されている移動端末毎に、前記音声認識手段による音声認識処理によって得られる認識音声データを、それに対応する移動端末に返信する認識音声データ返信手段と、

を含むことを特徴とする音声制御ホスト装置。

【請求項4】 前記移動端末内のホスト接続手段は、前記音声制御ホスト装置に対して、認識音声データをリアルタイム又は非リアルタイムに返信させるためのリアルタイム返信要求コマンド又は非リアルタイム返信要求コマンドを選択的に送信する機能と、認識音声データを一括して転送させるための一括転送要求コマンドを送信する機能とを含み、

前記認識音声データ返信手段は、現在接続されている移動端末毎に、前記移動端末接続手段が前記移動端末から前記リアルタイム返信要求コマンドを受信している場合には、それに対応して前記音声認識手段による音声認識処理によって得られる認識音声データをそれに対応する移動端末に即座に返信し、前記移動端末接続手段が前記移動端末から前記非リアルタイム返信要求コマンドを受信している場合には、それに対応して前記音声認識手段による音声認識処理によって得られる認識音声データを保持した後、前記移動端末接続手段が前記移動端末から前記一括転送要求コマンドを受信した時点で、それに対応して前記保持していた認識音声データをそれに対応する移動端末に一括して返信する、

ことを特徴とする請求項1乃至3の何れか1項に記載の移動端末音声認識通信システム、移動端末、又は音声制御ホスト装置。

【請求項5】 前記移動端末は、パーソナルハンディホンシステム通信機能を備え、

前記中継網は、パーソナルハンディホンシステム通信網とインターネットを含み、

前記音声制御ホスト装置は、前記インターネットに接続し、

前記移動端末内のホスト接続手段は、前記パーソナルハンディホンシステム通信網を介して、該パーソナルハンディホンシステム通信網を含む公衆網と前記インターネットとの間のゲートウェイ機能を有する移動端末制御ホスト装置に発信し接続することによって、前記インターネット上の通信プロトコルを使用して、前記移動端末制御ホスト装置から前記インターネットを介して前記音声制御ホスト装置に接続する、

ことを特徴とする請求項1乃至4の何れか1項に記載の

移動端末音声認識通信システム、移動端末、又は音声制御ホスト装置。

【請求項6】 前記通信プロトコルは、インターネットプロトコル層及びトランスミッションコントロールプロトコル層を含む階層プロトコルであって、

前記インターネット上を伝送される前記インターネットプロトコル層のパケットデータであるインターネットプロトコルデータグラムのヘッダフィールドには、前記インターネット上での前記移動端末及び前記音声制御ホスト装置のアドレスを指定する送信元インターネットプロトコルアドレス及び宛先インターネットプロトコルアドレスが格納され、該インターネットプロトコルデータグラムのデータフィールドには、前記トランスミッションコントロールプロトコル層のパケットデータであるトランスミッションコントロールプロトコルセグメントが格納され、

前記トランスミッションコントロールプロトコルセグメントのヘッダフィールドには、前記音声認識処理のための通信プロトコルを特定する送信元ポート番号及び宛先ポート番号が格納され、該トランスミッションコントロールプロトコルセグメントのデータフィールドには、前記移動端末を識別するための端末識別コード、前記リアルタイム返信要求コマンド、前記非リアルタイム返信要求コマンド、前記一括転送要求コマンド、前記音声データ、又は前記認識音声データが格納される、ことを特徴とする請求項5に記載の移動端末音声認識通信システム、移動端末、又は音声制御ホスト装置。

【請求項7】 前記音声制御ホスト装置は、網によって相互に接続され、前記移動端末接続手段、前記音声データ受信手段、前記音声認識手段、及び前記認識音声データ返信手段に対応する機能を分散して実現する複数のホストコンピュータから構成される、

ことを特徴とする請求項1乃至6の何れか1項に記載の移動端末音声認識通信システム、移動端末、又は音声制御ホスト装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、移動（携帯）端末装置において入力された通話音声等の音声を認識する技術に関する。

【0002】

【従来の技術及び発明が解決しようとする課題】 音声信号を認識して、文字データに変換して蓄積したり、認識結果を種々のサービスの利用に供したりするための音声認識技術は、様々な産業分野で従来から要請されている。

【0003】 近年では、音声認識アルゴリズムの発達により、メインフレームコンピュータ又はワークステーションコンピュータ等を用いた音声認識システムが開発されている。

【0004】 これらのシステムは、例えば、電話音声を入力とする銀行の残高照会システムや座席予約システム、作業員の音声を認識して荷物の自動配送を行う配送荷物の仕分システムを始めとして、種々の産業分野に取り入れつつある。

【0005】 しかし、このような音声認識システムは、上述のような大規模なコンピュータシステムの環境のもとでようやく実用的な認識精度を有するレベルに到達したばかりであり、いわゆるパーソナルコンピュータのような小型のコンピュータシステム環境のもとでは、実用的な認識精度を有する安価な音声認識システムは未だ実現されていないのが現状である。

【0006】 一方、上述のような情報処理技術と並行して、近年、自動車電話・携帯電話やPHS（パーソナルハンディホンシステム）等の移動端末が、急速に普及しつつある。

【0007】 特に、PHSは、小型であると同時に、自動車電話・携帯電話に比較して通話料金が安く、かつ、「いつでも、どこでも、誰とでも」高い品質で通話ができるという特徴を備えており、爆発的に普及しつつある。更に、PHSは、ISDN（Integrated Services Digital Network：サービス統合デジタル網）をバックボーンとする公衆網であるため、32キロビット/秒の伝送レートでの高速デジタル通信が可能であり、マルチメディア通信分野への応用に対する期待も高まっている。

【0008】 更には、移動端末の利便性をいかにすべく、携帯電話装置としてだけでなく、携帯情報管理装置としても利用できるように、マルチメディア情報管理/通信端末装置としての実現の可能性も期待されている。具体的には、このような移動端末は、通話機能/FAX機能を備えることはもちろん、インターネットや社内ネットワークへのアクセス機能としてのホームページアクセス機能や電子メール通信機能を備えることが予想されるほか、アドレス管理、スケジュール管理、データベース検索等の情報管理機能をも兼ね備えることが期待される。

【0009】 そして、このような移動端末は、人が気軽に利用できるよう、できる限り人にやさしく自然なユーザインタフェースを備えることが要請される。現在実現されているユーザインタフェースとしては、キーボードやマウスによる指操作入力、電子ペンによる手書き入力等が実用化されているが、音声入力等にも対応することができれば、ユーザインタフェースとして理想的である。即ち例えば、基本機能としての通話機能を利用しながら通話内容を示す音声信号をデータとして処理すること等が可能になれば、移動端末の利便性を飛躍的に増大させることが可能になる。ここに、移動端末に対してユーザインタフェースとして音声認識機能を適用することの価値を見出すことができる。

【0010】 しかし、移動端末は小型でありその情報処

理能力は限られている反面、前述したように、現在の音声認識処理では、メインフレームコンピュータ又はワークステーションコンピュータクラス的环境のもとでない、実用的な認識精度を実現することは困難である。従って、現状では、移動端末のユーザインタフェースとして音声認識機能を実現することは非常に困難である、という問題点を有している。

【0011】本発明の課題は、移動端末を用いた通信環境において、そのユーザインタフェースとしての音声認識機能を実用的な精度及びコストで実現することにある。

【0012】

【課題を解決するための手段】本発明はまず、以下の構成を含む移動端末を有する。即ち、ホスト接続手段（制御部110、通信部111）は、無線網又は有線網の何れか一方又は両方から構成される中継網（PHS網103とインターネット105）を介して間接的に又はその中継網を介さずに直接的にホスト装置である音声制御ホスト装置（音声制御ホスト装置108）に接続する。この手段は、例えば、音声制御ホスト装置に対して、認識音声データをリアルタイム又は非リアルタイムに返信させるためのリアルタイム返信要求コマンド又は非リアルタイム返信要求コマンドを選択的に送信する機能と、認識音声データを一括して転送させるための一括転送要求コマンドを送信する機能とを含む。

【0013】音声入力手段（入力部109）は、音声を入力する。音声データ送信手段（制御部110、通信部111）は、ホスト接続手段による接続動作の後、音声入力手段から入力される音声データを音声制御ホスト装置に送信する。

【0014】認識音声データ受信手段（制御部110、通信部111）は、音声制御ホスト装置から返信される認識音声データを受信する。認識音声データ表示/編集手段（制御部110、出力部112）は、その受信された認識音声データを表示又は編集する。

【0015】次に、本発明は、以下の構成を含む音声制御ホスト装置108を有する。即ち、移動端末接続手段（パケット送受信部115、移動端末通信制御部116）は、移動端末内のホスト接続手段による接続動作に応答して、移動端末を識別して接続する。

【0016】音声データ受信手段（パケット送受信部115、移動端末通信制御部116）は、現在接続されている移動端末毎に、音声データを受信する。音声認識手段（移動端末通信制御部116、文音声認識部117）は、現在接続されている移動端末毎に、音声データ受信手段により受信された音声データに対して音声認識処理を実行する。

【0017】認識音声データ返信手段（移動端末通信制御部116、パケット送受信部115）は、現在接続されている移動端末毎に、音声認識手段による音声認識処

理によって得られる認識音声データを、それに対応する移動端末に返信する。この手段は、例えば、現在接続されている移動端末毎に、移動端末接続手段が移動端末からリアルタイム返信要求コマンドを受信している場合には、それに対応して音声認識手段による音声認識処理によって得られる認識音声データをそれに対応する移動端末に即座に返信し、移動端末接続手段が移動端末から非リアルタイム返信要求コマンドを受信している場合には、それに対応して音声認識手段による音声認識処理によって得られる認識音声データを保持した後、移動端末接続手段が移動端末から一括転送要求コマンドを受信した時点で、それに対応して保持していた認識音声データをそれに対応する移動端末に一括して返信する。

【0018】以上の移動端末と音声制御ホスト装置を含む本発明による通信移動端末音声認識通信システムにより、移動端末は、高度な音声認識環境を設備する必要がなく実用的な精度を有する音声認識機能の提供を低コストで受けることができる。

【0019】上述の発明の構成は、下記の限定を含むことができる。即ち、まず、移動端末は、パーソナルハンディホンシステム通信機能（通信部111）を備える。

【0020】次に、中継網は、パーソナルハンディホンシステム通信網（PHS網103）とインターネット（インターネット105）を含む。また、音声制御ホスト装置は、インターネットに接続する。

【0021】そして、移動端末内のホスト接続手段は、パーソナルハンディホンシステム通信網を介して、そのパーソナルハンディホンシステム通信網を含む公衆網とインターネットとの間のゲートウェイ機能を有する移動端末制御ホスト装置（移動端末制御ホスト装置104）に発信し接続することによって、インターネット上の通信プロトコルを使用して、移動端末制御ホスト装置からインターネットを介して音声制御ホスト装置に接続する。

【0022】この限定構成によって、現在全国的及び全世界的に普及しつつあるパーソナルハンディホンシステム通信網及びインターネットを経由することにより、実用的な精度を有する音声認識機能の提供をより低コスト及び手軽に受けることができると同時に、本発明が提供する機能とパーソナルハンディホンシステム通話機能及びインターネットアクセス機能とを、シームレスに結合することができる。

【0023】更に、上述の通信プロトコルは、下記の限定を含むことができる。即ち、上記通信プロトコルは、インターネットプロトコル（IP）層及びトランスミッションコントロールプロトコル（TCP）層を含む階層プロトコルである。

【0024】次に、インターネット上を伝送されるインターネットプロトコル層のパケットデータであるインターネットプロトコルデータグラム（IPデータグラム）

10

20

30

40

50

のヘッダ(IPヘッダ)フィールドには、インターネット上での移動端末及び音声制御ホスト装置のアドレスを指定する送信元インターネットプロトコルアドレス及び宛先インターネットプロトコルアドレスが格納され、そのインターネットプロトコルデータグラムのデータフィールドには、トランスミッションコントロールプロトコル層のパケットデータであるトランスミッションコントロールプロトコルセグメントが格納される。

【0025】また、トランスミッションコントロールプロトコルセグメント(TCPセグメント)のヘッダ(TCPヘッダ)フィールドには、音声認識処理のための通信プロトコルを特定する送信元ポート番号及び宛先ポート番号が格納され、そのトランスミッションコントロールプロトコルセグメントのデータフィールドには、移動端末を識別するための端末識別コード、リアルタイム返信要求コマンド、非リアルタイム返信要求コマンド、一括転送要求コマンド、音声データ、又は認識音声データが格納される。

【0026】この限定構成によって、移動端末と音声制御ホスト装置とを全世界的に容易に特定できると共に、音声認識処理サービスと他の情報処理サービスとの共存を容易に実現できる。

【0027】ここまでの発明の構成において、音声制御ホスト装置は、網によって相互に接続され、移動端末接続手段、音声データ受信手段、音声認識手段、及び認識音声データ返信手段に対応する機能を分散して実現する複数のホストコンピュータから構成されるように実現することができる。

【0028】この限定構成によって、ホスト装置側の負荷分散を容易に実現できる。なお、上述した移動端末及び音声制御ホスト装置の単体も、本発明の権利範囲である。

【0029】

【発明の実施の形態】以下、図面を参照しながら本発明の実施の形態について詳細に説明する。本実施の形態では、PHS機能が組み込まれた移動端末において、通話時に又はオフライン状態でマイクから入力された音声信号が、PHS網からインターネットを介して特定の音声サービスプロバイダ内のLANに接続される音声制御ホスト装置に送られ、そこで上記音声信号が認識された後、リアルタイムに又は所望のタイミングで一括して移動端末に返送され、移動端末での情報管理処理に利用されることが、本発明に関連する大きな特徴である。このようなシステムにより、移動端末は、高度な音声認識環境を設備する必要がなく実用的な精度を有する音声認識機能の提供を低コストで受けることができる。

<システム構成>図1は、本発明の実施の形態の全体システム構成図である。

【0030】移動端末101は、PHS端末機能を有しており、無線基地102を介して、無線通信によってP

HS網103に接続される。無線基地102は、街路の公衆電話ボックス、電柱、ビル屋上、地下通路等に設けられる公衆無線基地、又は加入者宅内の親子電話装置等である。なお、親子電話装置に接続される場合は、PHS網を介さずに、直接公衆電話網に接続される。なお、無線基地102の代わりに、有線接続装置を介して、有線通信によってPHS網103又は公衆電話網に接続されるように構成されてもよい。

【0031】PHS網103は、公衆電話網又はISDN網と相互接続しており、これらの網には、高速デジタル専用線等によってインターネット105に接続している移動端末制御ホスト装置104が接続されている。

【0032】移動端末101は、無線基地102及びPHS網103を介して、上記公衆電話網又はISDN網に接続されている移動端末制御ホスト装置104に自動的にダイヤルアップ発信することによって、インターネット105に接続することができる。

【0033】インターネット105には、高速デジタル専用線等を介して所定の音声サービスプロバイダのLAN107に接続しているルータ装置106が接続されている。LAN107は、イーサネット方式、ATM(Asynchronous Transfer Mode)方式、又はFDDI方式によるローカルエリアネットワークである。LAN107には、更に音声制御ホスト装置108が接続されている。

【0034】移動端末101は、移動端末制御ホスト装置104に自動的にダイヤルアップ発信した後に、インターネット105、ルータ装置106、及びLAN107を介して、音声制御ホスト装置108と通信することができる。

【0035】今、移動端末101内の入力部109において、ユーザが、タッチパネルから音声制御ホスト装置108との通信を指示すると、制御部110は、通信部111に対して、音声制御ホスト装置108との通信開始を依頼する。

【0036】通信部111は、制御部110から通信開始を依頼されると、現在移動端末制御ホスト装置104に接続していなければ、無線基地(又は有線接続装置)102に無線(又は有線)発信してPHS網103に接続した後、移動端末制御ホスト装置104のアクセス電話番号を指定してダイヤルアップ発信する。

【0037】移動端末制御ホスト装置104が着信すると、移動端末101内の通信部111は、まず、移動端末制御ホスト装置104内の接続確立部113と通信することにより、インターネット105上の標準通信プロトコルであるTCP/IP及びPPP方式による接続の確立のためのネゴシエーションを行う。この結果、移動端末制御ホスト装置104から、移動端末101内の通信部111に対して、インターネット105上の識別アドレスであるIPアドレスが付与され、移動端末101

は、インターネット105へのアクセスが可能となる。

【0038】移動端末101内の通信部111は、既に移動端末制御ホスト装置104に接続していれば、上記ダイヤルアップ発信は省略する。その後、移動端末101内の通信部111は、予め設定されている音声制御ホスト装置108のIPアドレスである“宛先IPアドレス”と、移動端末制御ホスト装置104から付与されたIPアドレスである“送信元IPアドレス”と、移動端末101を識別するための“端末識別コード”（例えばPHS電話番号）と、ユーザの指定に基づく文音声認識処理のリアルタイム開始要求コマンド又は文音声認識処理の非リアルタイム開始要求コマンドとが格納されたTCP/IPパケットを、インターネット105に向けて送出する。

【0039】このTCP/IPパケットは、それに格納されている“宛先IPアドレス”に基づき、移動端末制御ホスト装置104内のルーティング部114とインターネット105内の特には図示しない中継ホスト装置によって、音声サービスプロバイダ内のルータ装置106まで転送された後、更に、LAN107を介して音声制御ホスト装置108内のパケット送受信部115まで転送される。

【0040】パケット送受信部115は、受信したTCP/IPパケットから、“送信元IPアドレス”と、“端末識別コード”と、文音声認識処理のリアルタイム開始要求コマンド又は文音声認識処理の非リアルタイム開始要求コマンドとを取り出して、音声制御ホスト装置108内の移動端末通信制御部116に引き渡す。

【0041】移動端末通信制御部116は、引き渡された“送信元IPアドレス”と、“端末識別コード”と、文音声認識処理のリアルタイム開始要求コマンド又は文音声認識処理の非リアルタイム開始要求コマンドに関する情報を後述する処理端末登録テーブル（図13）に登録した後、パケット送受信部115に対して、送信許可データが格納されたTCP/IPパケットの移動端末101への返信を依頼する。

【0042】パケット送受信部115は、対応するTCP/IPパケットを、移動端末101に対応するIPアドレスに向けて送信する。このようにして、音声制御ホスト装置108は、移動端末101から転送されてくる音声データに対して文音声認識処理を実行することが可能となる。

【0043】移動端末101内の通信部111は、音声制御ホスト装置108から上記送信許可データが格納されたTCP/IPパケットを受信すると、それに格納されている送信許可データを制御部110に引き渡す。

【0044】移動端末101内の制御部110は、上記送信許可データを引き渡された後、通信部111に対して、通話動作又はオフライン状態での音声入力動作によってマイクから入力された音声データの音声制御ホスト

装置108への送信を依頼する。

【0045】通信部111は、上記音声データが格納されたTCP/IPパケットを、音声制御ホスト装置108に対応するIPアドレスに向けて送信する。このTCP/IPパケットは、それに格納されている“宛先IPアドレス”に基づき、移動端末制御ホスト装置104内のルーティング部114、インターネット105内の特には図示しない中継ホスト装置、音声サービスプロバイダ内のルータ装置106、及びLAN107を介して、音声制御ホスト装置108内のパケット送受信部115まで転送される。

【0046】パケット送受信部115は、受信したTCP/IPパケットに格納されている音声データを取り出し、それを音声制御ホスト装置108内の移動端末通信制御部116に引き渡す。

【0047】移動端末通信制御部116は、引き渡された音声データを文音声認識部117に引き渡す。文音声認識部117は、引き渡された音声データに対して文音声認識処理を実行し、認識結果である認識音声文章データを移動端末通信制御部116に引き渡す。

【0048】移動端末通信制御部116は、移動端末101から文音声認識処理のリアルタイム開始要求コマンドが指定されている場合には、即座に、パケット送受信部115に対して、認識音声文章データが格納されたTCP/IPパケットの移動端末101への返信を依頼する。

【0049】パケット送受信部115は、対応するTCP/IPパケットを、移動端末101に対応するIPアドレスに向けて送信する。移動端末101内の通信部111は、音声制御ホスト装置108から上記認識音声文章データが格納されたTCP/IPパケットを受信すると、それに格納されている認識音声文章データを制御部110に引き渡す。

【0050】移動端末101内の制御部110は、上記認識音声文章データを、出力部112に出力する。出力部112は、認識音声文章データに対応する文章を、LCD表示部に表示する。ユーザは、この文章データを、任意に蓄積又は加工することができる。

【0051】移動端末通信制御部116は、移動端末101から文音声認識処理の非リアルタイム開始要求コマンドが指定されている場合には、認識音声文章データを即座に移動端末101に返信することはせずに、それを順次蓄積する。そして、移動端末通信制御部116は、後に、移動端末101から文音声認識結果の一括転送要求コマンドを受信したタイミングで、パケット送受信部115に対して、蓄積しておいた認識音声文章データが格納されたTCP/IPパケットの移動端末101への一括返信を依頼する。

【0052】パケット送受信部115は、対応するTCP/IPパケットを、移動端末101に対応するIPア

10

20

30

40

50



ドレスに向けて送信する。これにより、移動端末101では、一括転送された認識音声文章データを、LCD表示部に表示させ、ユーザは、この文章データを、任意に蓄積又は加工することができる。

【0053】一方、移動端末101は、音声制御ホスト装置108との通信のほかに、それが装備するホームページ閲覧ツールや電子メールを利用して、移動端末制御ホスト装置104にダイヤルアップ発信することにより、インターネット105上の所望のリソースに自由にアクセスすることが可能である。＜移動端末101の外観構成＞図2は、図1の移動端末101の外観図である。

【0054】移動端末101は、コンパクトな携帯情報管理装置の外観を有し、音声を入力するための送話器を兼ねたマイク201と、本発明には特には関連しないが画像を入力するためのカメラ202と、各種情報を表示し、またタッチ入力又はペン入力を受け付けるタッチパネル機能を有するLCD表示部203と、音声を出力するための受話器を兼ねたスピーカ204を有する、また、図1の無線基地102に発信するための無線アンテナ205と、無線基地102の代わりに有線接続装置に接続するためのソケット206を有する。

【0055】更に、各種ICカードを挿入するためのICカードスロット207と、他の移動端末101又はパーソナルコンピュータ等との間で赤外線光通信を行うための光送受信機208を有する。

【0056】スイッチ209は、電源スイッチである。＜移動端末101の機能ブロック構成＞図3は、移動端末101の機能ブロック図である。

【0057】移動端末101は、図1にも示したように、入力部109、制御部110、通信部111、及び出力部112から構成され、それぞれバス326によって相互に接続されている。

【0058】まず、入力部109は、音声を入力する部分と、本発明には特には関連しないが画像を入力する部分と、出力部112の動作において後述するタッチパネル機構の部分とから構成される。

【0059】音声を入力する部分は、マイク301、A/D変換部302、及びマイク制御部303から構成される。マイク301（図2の201に対応）は、PHS電話の送話器を兼ねており、ユーザが発声した音声を入力する。

【0060】A/D変換部302は、マイク301から入力されたアナログ音声信号をデジタル音声データに変換し、更にそのデジタル音声データを、PHSの標準音声符号化方式であるADPCM（Adaptive Differential Pulse Code Modulation: 適応差分線形パルス符号化）方式によって符号化する。なお、この部分は、PHS端末を構成するLSI集積回路として、既に実用化されている。

【0061】マイク制御部303は、上述の符号化された音声データを、通話時には、通信部111内の通信制御部321に転送して通話チャネルに載せると共に、文音声認識処理時には、更に制御部110内のRAM317に転送する。

【0062】一方、画像を入力する部分は、CCD（Charge Coupled Device）カメラ304、A/D変換部305、メモリ306、及びカメラ制御部307から構成される。

【0063】CCDカメラ304は、ユーザの操作に基づいて任意の画像を撮像する。A/D変換部305は、CCDカメラ304によって撮像されたアナログ映像信号を、デジタル画像データに変換する。

【0064】メモリ306は、デジタル画像データをフレーム単位で記憶する。カメラ制御部307は、CCDカメラ304、A/D変換部305、及びメモリ306の動作を制御する。

【0065】次に、出力部112は、音声を出力する部分と、画像を出力する部分とから構成される。音声を出力する部分は、スピーカ308、D/A変換部309、及びスピーカ制御部310から構成される。

【0066】スピーカ制御部310は、通信部111内の通信制御部321から受信されたPHS通話音声データ、又は制御部110内のRAM317から受信された合成音声データを、D/A変換部309に転送する。

【0067】D/A変換部309は、受信された音声データを復号し、アナログ音声信号に変換し、それをスピーカ308（図2の204に対応）から音声として放音させる。

【0068】画像を出力する部分は、LCD表示部203、LCDドライバ312、メモリ313、及びLCD制御部314から構成される。LCD制御部314は、制御部110内のRAM317から受信された文字データ、イメージデータ、コマンドボタンデータ等の各種画像データをメモリ313にフレーム単位で保持させ、LCDドライバ312に起動をかける。

【0069】LCDドライバ312は、メモリ313からフレーム単位で読み出される画像データを、LCD表示部311（図2の203に対応）に表示する。なお、LCD表示部311（図2の203）の表面には、透明タッチパネルが配設されており、ユーザは、LCD表示部311に表示されるコマンドボタンデータ等に従って、タッチパネルに指タッチ又はペンタッチすることにより、コマンド入力を行うことができる。この入力信号は、タッチパネル制御部315によって制御部110内のRAM317に転送される。

【0070】続いて、制御部110は、CPU316、RAM317、及びROM318と、ICカードインタフェース部319、及び必要に応じてICカードスロット207（図2）に挿入されるICカード320とから



構成される。

【0071】CPU316は、ROM318に記憶された制御プログラムに従って、RAM317をワークエリアとして使用しながら、移動端末101全体の動作を制御する。

【0072】ICカードインタフェース部319は、ICカード320に対するデータの入出力を制御する。最後に、通信部111は、通信制御部321、無線ドライバ322、無線アンテナ323、有線ドライバ324、及びソケット325から構成される。

【0073】通信制御部321は、PHS通話処理及びインターネット105との間のTCP/IP通信処理（後述する）を実行し、無線ドライバ322又は有線ドライバ324を制御する。

【0074】無線ドライバ322は、無線通信時に、通信データを、無線アンテナ323（図2の205に対応）を介して送受信されるPHS無線信号との間で相互変換する。PHS無線信号は、1.9GHzの無線周波数と、300kHzのキャリア周波数間隔と、4チャネル／キャリアのTDMA-TDD無線アクセス方式と、 $\pi/4$ シフトQPSK変調方式と、384kbits/secの無線伝送速度に基づく無線信号である。

【0075】一方、有線ドライバ324は、有線通信時に、通信データを、ソケット325（図2の206に対応）を介して送受信される有線信号との間で相互変換する。これは、一般的な電話帯域モデム変調信号である。以上の構成を有する本発明の実施の形態の動作について、以下に詳細に説明する。

<移動端末101の処理>まず、移動端末101の処理について説明する。

【0076】図4は、図3の制御部110内のCPU316が、電源投入後に、制御部110内のROM318に記憶されている制御プログラムを実行する動作として実現される制御動作を示す全体動作フローチャートである。

【0077】なお、図4、図5、及び図8の動作フローチャートで示される各機能を実現する制御プログラム及びそれに必要なデータは、例えば、図2に示されるICカードスロット207に着脱自在なICカード320に、CPU316が読み取り可能なプログラムコードの形態で記憶され、そのプログラムコードがCPU316によって直接実行され、又は、そのプログラムコードが必要に応じてRAM317又は書込み可能なROM318にロードされてCPU316によって実行されるように構成されてもよい。或いは、上述の制御プログラム及びそれに必要なデータは、無線又は有線の通信回線又は光送受信機208（図2）から通信部111を介して他の機器から受信されて、RAM317又は書込み可能なROM318にロードされてCPU316によって実行されるように構成されてもよい。

【0078】まず、ステップ401→402→403→404→401の繰返しループにおいては、図3のタッチパネル制御部315からタッチパネル入力検出が通知されたか否かの判定処理（401）、音声制御ホスト装置108（図1）から認識音声文章データが受信されたか否かの判定処理（402）、その他の受信／表示処理（403）、及び必要なデータの送信処理（404）が実行される。

【0079】タッチパネル制御部315からタッチパネル入力検出が通知されステップ401の判定がYESとなると、ステップ405又は406で、上記タッチパネル入力検出が図3のCCDカメラ304（図2の202）の入力指示又は図3のマイク301（図2の201）の入力指示であるか否かが、判定される。

【0080】タッチパネル入力検出が図3のCCDカメラ304（図2の202）の入力指示であってステップ405の判定がYESとなると、ステップ407で、図3の入力部109内のカメラ制御部307に対して、例えば手書き文字画像等の入力処理の開始が指示される。その後、ステップ404の送信処理に進む。画像入力処理は、本発明には特に関連しないため、その詳細な説明は省略する。

【0081】タッチパネル入力検出が図3のマイク301（図2の201）の入力指示であってステップ406の判定がYESとなると、ステップ408で、図3の入力部109内のマイク制御部303に対して、音声入力処理の開始が指示される。この音声入力処理の開始指示は、例えばPHS通話処理の開始指示、又は文音声認識処理を実行するためのオフライン状態での音声入力処理の開始指示である。

【0082】マイク制御部303は、上述のCPU316からの指示によって、マイク301（図2の201）及びA/D変換部302に対して、音声入力処理の開始を指示する。この結果、A/D変換部302からは、マイク301（図2の201）から入力された音声データが出力される。

【0083】その後、上述の音声入力処理の開始指示がPHS通話の開始指示である場合には、上述の音声データは、通信制御部321の特には図示しない送信処理によって、所定の通話チャネルに載せられて通話相手に送信される。

【0084】また、上述の音声入力処理の開始指示が文音声認識処理のための音声入力処理の開始指示を含む場合には、それ以後マイク301（図2の201）から入力されマイク制御部303から出力された音声データは、後述するステップ404の送信処理において、そこで音声制御ホスト装置108に向けて送信される。

【0085】タッチパネル入力検出が図3のCCDカメラ304（図2の202）の入力指示でも図3のマイク301（図2の201）の入力指示でもない場合には、ステ

ップ405及び406の判定がNOとなって、ステップ409で、他のキー入力処理が実行される。その後、ステップ404の送信処理に進む。

【0086】一方、音声制御ホスト装置108（図1）から通信部111を介して制御部110内のRAM317に認識音声文章データが受信され、ステップ401→402→403→404→401の繰返しループにおけるステップ402の判定がYESとなると、ステップ410において、上記認識音声文章データがRAM317から出力部112内のメモリ313に転送され、LCD制御部314に対して上記認識音声文章データの表示が指示される。

【0087】この結果、LCD制御部314の制御によって、メモリ313からLCDドライバ312を介してLCD表示部311（図2の203）に、受信された認識音声文章データが表示される。

【0088】次に、ステップ404の送信処理について説明する。図5は、上記送信処理の詳細を示す動作フローチャートである。まず、ステップ501では、図4のステップ409の他キー入力処理によって処理されたタッチパネルからのキー入力送信指示を伴っているか否かが判定される。

【0089】この判定がNOの場合には、ステップ505の処理へ進む。ステップ501の判定がYESの場合には、ステップ502で、移動端末101が現在図1の移動端末制御ホスト装置104に接続中であるか否かが判定される。

【0090】移動端末101が現在図1の移動端末制御ホスト装置104に接続中でありステップ502の判定がYESならば、図3の制御部110内のCPU316は、ステップ504で、移動端末101の“端末識別コード”とキー入力処理に対応するコマンドの送信指示を、図3の通信部111内の通信制御部321に対し依頼する。この結果、通信制御部321は、上記“端末識別コード”とコマンドが格納されたTCP/IPパケットを生成し、それをインターネット105に接続されている所定のホスト（例えば図1の音声制御ホスト装置108）に向け送信する。

【0091】移動端末101が現在図1の移動端末制御ホスト装置104に接続中ではなくステップ502の判定がNOならば、図3の制御部110内のCPU316は、ステップ503で、図3の通信部111内の通信制御部321に対して発信処理を依頼してから、ステップ504を実行する。

【0092】後に詳述するように、ユーザの指定に基づく文音声認識処理のリアルタイム開始要求コマンド又は文音声認識処理の非リアルタイム開始要求コマンドの送信指示、文音声認識処理の終了要求コマンドの送信指示、及び文音声認識処理の一括転送要求コマンドの送信指示は、上述のステップ504において発行される。

【0093】前述したようにステップ501の判定がNOの場合又はステップ504の処理の後、ステップ505では、図4のステップ408によって、文音声認識処理のための音声入力処理の開始指示が実行されており、音声データの音声制御ホスト装置108（図1）への送信指示がなされているか否かが判定される。

【0094】この判定がNOの場合には、ステップ510の処理へ進む。ステップ505の判定がYESの場合には、ステップ506で、音声制御ホスト装置108から文音声認識処理のリアルタイム開始要求コマンド又は文音声認識処理の非リアルタイム開始要求コマンドに対する応答である送信許可データが既に返信されているか否かが判定される。

【0095】この判定がNOの場合には、音声制御ホスト装置108がまだ移動端末101からの文音声認識処理のリアルタイム開始要求コマンド又は文音声認識処理の非リアルタイム開始要求コマンドに対する準備が完了していないため、ステップ510の処理へ進む。

【0096】音声制御ホスト装置108から文音声認識処理のリアルタイム開始要求コマンド又は文音声認識処理の非リアルタイム開始要求コマンドに対する応答である送信許可データが既に返信されておりステップ506の判定がYESの場合には、更に、ステップ507で、移動端末101が現在図1の移動端末制御ホスト装置104に接続中であるか否かが判定される。

【0097】移動端末101が現在図1の移動端末制御ホスト装置104に接続中でありステップ507の判定がYESならば、図3の制御部110内のCPU316は、ステップ509で、図3に示される入力部109内のマイク制御部303から制御部110内のRAM317に転送されてきている音声データの送信指示を、通信部111内の通信制御部321に対し依頼する。この結果、通信制御部321は、上記音声データが格納されたTCP/IPパケットを生成し、それをインターネット105に接続されている図1の音声制御ホスト装置108に向けて送信する。

【0098】移動端末101が現在図1の移動端末制御ホスト装置104に接続中ではなくステップ507の判定がNOならば、図3の制御部110内のCPU316は、ステップ508で、図3の通信部111内の通信制御部321に対して発信処理を依頼してから、ステップ509を実行する。

【0099】後に詳述するように、文音声認識処理のための音声データの送信指示は、上述のステップ509において発行される。前述したようにステップ505又は506の判定がNOの場合又はステップ509の処理の後、ステップ510では、図4のステップ407によって、画像入力処理の開始指示が実行されており、画像データを図1のインターネット105に接続されている特には図示しない画像制御ホスト装置への送信指示がなさ

れているか否かが判定される。

【0100】この判定がNOの場合には、図4のステップ404の送信処理を終了する。ステップ510の判定がYESの場合には、ステップ511で、移動端末101が現在図1の移動端末制御ホスト装置104に接続中であるか否かが判定される。

【0101】移動端末101が現在図1の移動端末制御ホスト装置104に接続中でありステップ511の判定がYESならば、図3の制御部110内のCPU316は、ステップ513で、図3に示される入力部109内のメモリ306に得られている画像データの送信指示を、通信部111内の通信制御部321に対して依頼する。この結果、通信制御部321は、上記画像データが格納されたTCP/IPパケットを生成し、それをインターネット105に接続されている特には図示しない画像制御ホスト装置108に向けて送信する。

【0102】移動端末101が現在図1の移動端末制御ホスト装置104に接続中ではなくステップ511の判定がNOならば、図3の制御部110内のCPU316は、ステップ512で、図3の通信部111内の通信制御部321に対して発信処理を依頼してから、ステップ513を実行する。

【0103】なお、ステップ513の画像データの送信指示は、本発明には特には関連しないため、その詳細な説明は省略する。前述したようにステップ510の判定がNOの場合又はステップ513の処理の後、図4のステップ404の送信処理を終了する。

<通信データのフォーマット>図6は、移動端末101と移動端末制御ホスト装置104及びインターネット105（音声制御ホスト装置108）との間で通信される通信データのフォーマット図である。

【0104】移動端末101と移動端末制御ホスト装置104との間では、通信データは、PPP（Point-to-Point Protocol）と呼ばれる通信プロトコルに基づき、図6(a)に示されるPPPフレーム（図の左から右に向けて転送される）を用いて、PHS規格の32kbit/s/secの伝送レートを有するデジタル通信チャネル上を伝送される。

【0105】PPPフレームを構成する、“フラグ”、“アドレス”、“コントロール”の各フィールドは、図6(a)に示される各固定ビット列が設定される。2オクテットのデータ長を有するFCSは、フレームチェックシーケンスと呼ばれ、PPPフレームデータの誤り検出／訂正用のデータである。移動端末101と移動端末制御ホスト装置104との間でPPPリンクが確立した後には転送されるPPPフレームの“インフォメーション”フィールド（可変データ長を有する）には、インターネット105（図1）上のデータの基本伝送単位であるIPデータグラムが格納され、その場合に、2オクテットのデータ長を有する“プロトコル”フィールドには、“

インフォメーション”フィールドにIPデータグラムが格納されていることを示す16進値“0021”が格納される。

【0106】PPPフレームの“インフォメーション”フィールドには、上述のようにIPデータグラムが格納される。このIPデータグラムは、上述のようにインターネット105上のデータの基本伝送単位である。IPデータグラムは、インターネットプロトコル（IP）に従って規定され、その“データ”フィールドに格納されたデータをインターネット105上の宛先のホスト装置まで一意に転送するための機能を提供し、インターネット105上でのアドレスを特定する機能、そのIPデータグラム自身を“宛先IPアドレス”で指定されたホストまでインターネット105上の一定の経路で転送する機能、そのIPデータグラム自身のフラグメント化（分割）と再組立てを行う機能等を備える。

【0107】IPデータグラムは、図6(b)に示されるように、IPヘッダフィールドとデータフィールドとから構成される。IPヘッダフィールドには、それが含まれるIPデータグラム自身を配送するために必要な全ての情報が含まれる。図7(a)は、IPヘッダのフォーマット図である。

【0108】IPヘッダは、32ビットを1ワードとして、5乃至6ワードのデータ長を有し、このデータ長は第1ワードの“ヘッダ長”フィールドに格納され、また、IPデータグラム全体のデータ長は、第1ワードの“IPデータグラムの全長”フィールドに格納される。

【0109】第1ワードの“バージョン”フィールドには、IPデータグラムの転送方法を規定するインターネットプロトコル（IP）のバージョンが設定され、現在のバージョンは4である。

【0110】第1ワードの“サービスの種類”フィールドには、配送の優先度を表わす情報等が格納されるが、ここは本発明には特には関連しない。第2ワードの各フィールドは、IPデータグラムがインターネット105上での転送の制約によりフラグメント化（分割）される場合における制御情報を規定する。まず、“識別番号”フィールドには、分割されたフラグメントであるこのIPデータグラムが属する分割前のIPデータグラムを識別するための一意な整数が設定される。次に、“フラグメントのオフセット”フィールドには、分割されたフラグメントであるこのIPデータグラムが分割前のIPデータグラムのどの部分に相当するかを示すオフセット情報が設定される。そして、“フラグ列”フィールドには、分割されたフラグメントであるこのIPデータグラムに、それが属する分割前のIPデータグラムを構成する他のフラグメントが後続するか否かが設定される。以上の情報により、インターネット105上の中継ホストにおいてIPデータグラムがフラグメント化されても、受信側で分割前のIPデータグラムを正確に復元するこ

とができる。

【0111】第3ワードの“生存期間”(TTL:Time To Live)フィールドには、そのIPデータグラムがインターネット105上にどれだけの時間の間存在することを許すかを示す秒単位の時間情報が設定される。インターネット105上の中継ホストは、IPデータグラムを処理する毎に上記フィールド値を減算し、値が0以下になったIPデータグラムはインターネット105上から廃棄する。これにより、インターネット105上での過度なトラフィックの発生が抑制される。なお、廃棄されたIPデータグラムに対する再送制御は、そのIPデータグラムに格納されるTCPセグメントに対する制御処理において実現される。

【0112】第3ワードの“プロトコル”フィールドには、そのIPデータグラムの“データ”フィールドに格納されるデータのフォーマットを規定するための整数値が設定される。本実施の形態の場合には、図6(c)に示されるように、IPデータグラムの“データ”フィールドにはTCPセグメントデータが格納されるため、そのフォーマットを規定する整数値6が設定される。

【0113】第3ワードの“ヘッダのチェックサム”フィールドには、IPヘッダのデータの誤りを検出するためのチェックサムデータが設定される。第4ワードには、32ビットの“送信元IPアドレス”が設定される。例えばIPデータグラムが移動端末101から音声制御ホスト装置108へ転送される場合には、“送信元IPアドレス”としては、後述する発信処理により移動端末制御ホスト装置104から移動端末101に対して付与されたIPアドレスが設定される。図1の音声制御ホスト装置108は、この“送信元IPアドレス”を記憶することにより、インターネット105を介して移動端末101に対して、フォーマット文章データ等を返信することができる。

【0114】第5ワードには、32ビットの“宛先IPアドレス”が設定される。例えばIPデータグラムが移動端末101から音声制御ホスト装置108へ転送される場合には、“宛先IPアドレス”としては、音声制御ホスト装置108に固定的に割当てられているIPアドレスが設定される。移動端末制御ホスト装置104内のルーティング部114、インターネット105上の各中継ホスト装置、及び音声サービスプロバイダ内のルータ装置106は、受信したIPデータグラムに格納されている上記“宛先IPアドレス”を識別することによって、予め各装置が有する経路制御テーブル情報に従って、そのIPデータグラムの配送経路を決定し、最終的にそのIPデータグラムを音声サービスプロバイダ内の音声制御ホスト装置108まで転送することができる。

【0115】第6ワードの“IPオプション”フィールドは、オプションであり、インターネット105を構成する各ネットワークのテスト又はデバッグのための情報

や、インターネット105上での配送経路を制御又は監視するための制御情報等が設定されるが、ここは本発明には特に関連しない。

【0116】第6ワードの“パディング”フィールドには、データ長を合わせるためのパディングデータが設定される。次に、IPデータグラムの“データ”フィールドには、TCPセグメントデータが格納される。このTCPセグメントは、トランスミッションコントロールプロトコル(TCP)に従って規定され、その“データ”フィールドに格納されたデータをインターネット105上の宛先のホスト装置まで正確に適切な順序で配送するための機能を備える。IPデータグラムがインターネット105上でのデータの一意な転送の機能のみを提供し、データの信頼性を確保する機能(再送制御機能等)を提供しないのに対して、TCPセグメントは、データの信頼性を確保する機能を提供するものである。

【0117】このように、通信データが、(PPPフレームと)IPデータグラムとTCPセグメントという階層構造を有するのは、インターネット105上ではなるべく小さい処理負荷のもとで効率良くデータを配送する必要があり、エンド対エンド間ではできるかぎり信頼性の高いデータ配送を実現する必要があるという異なる要請に効率的に対処するためである。これにより、インターネット105上の中継ホスト装置は、IPデータグラムのIPヘッダのみを参照することにより、そのIPデータグラムの“データ”フィールドに格納された情報(TCPセグメント)をできる限り高速かつ効率的に宛先ホスト装置まで配送することができ、エンド対エンド(送信元ホスト装置と宛先ホスト装置)間では、TCPセグメントのTCPヘッダを参照することにより、再送制御等の信頼性の高いデータ通信を実現することができるのである。

【0118】TCPセグメントは、図6(b)に示されるように、TCPヘッダフィールドとデータフィールドとから構成される。図7(b)は、TCPヘッダのフォーマット図である。

【0119】TCPヘッダは、IPヘッダの場合と同様に、32ビットを1ワードとして、5乃至6ワードのデータ長を有し、このデータ長は第4ワードの“ヘッダ長”フィールドに格納され、また、IPデータグラム全体のデータ長は、第1ワードの“IPデータグラムの全長”フィールドに格納される。

【0120】第1ワードの“送信元ポート番号”フィールドと“宛先ポート番号”フィールドには、文音声認識処理のための通信プロトコルを特定する16ビットの整数値が設定される。

【0121】音声制御ホスト装置108内のパケット送受信部115(図1)は、文音声認識処理のための音声データが格納されたTCPセグメントのほかにも、電子メールデータを始めとする様々なデータが格納された様

10

20

30

40

50

々なTCPセグメントを送受信するため、受信したTCPセグメントのTCPヘッダに設定されている“宛先ポート番号”フィールドの値を認識することによって、そのTCPセグメントの“データ”フィールドに格納されているデータを音声制御ホスト装置108で実行されるどのアプリケーションに引き渡すかを決定することができる。

【0122】そして、パケット送受信部115は、受信したTCPセグメントのTCPヘッダに設定されている“宛先ポート番号”フィールドの値が文音声認識処理のための通信プロトコルに対応する値を示している場合には、そのTCPセグメントの“データ”フィールドに格納されている音声データを移動端末通信制御部116に引き渡すことができる。

【0123】同様に、移動端末101の通信部111内の通信制御部321（図3）も、認識音声文章データが格納されたTCPセグメントのほかにも、ホームページデータや電子メールデータを始めとする様々なデータが格納された様々なTCPセグメントを送受信するため、受信したTCPセグメントのTCPヘッダに設定されている“宛先ポート番号”フィールドの値を認識することによって、そのTCPセグメントの“データ”フィールドに格納されているデータを移動端末101で実行されるどのアプリケーションに引き渡すかを決定することができる。

【0124】そして、通信制御部321は、受信したTCPセグメントのTCPヘッダに設定されている“宛先ポート番号”フィールドの値が文音声認識処理のための通信プロトコルに対応する値を示している場合には、制御部110（図1、図3）に、文音声認識処理のためのデータの受信を通知し、そのTCPセグメントの“データ”フィールドに格納されている認識音声文章データ等を引き渡すことができる。

【0125】更に、音声制御ホスト装置108内のパケット送受信部115及び移動端末101の通信部111内の通信制御部321は、受信したTCPセグメントのTCPヘッダに設定されている“送信元ポート番号”を確認することにより、送信元のアプリケーションを確認することができる。

【0126】次に、図7に示されるTCPヘッダの第2ワードの“シーケンス番号”フィールドは、現在のTCPコネクションにおいて送信側から受信側に送信される全バイトストリームのうち、このTCPセグメントの“データ”フィールドに格納されているデータの先頭が上記全バイトストリームの何バイト目にあたるかを、送信側から受信側に通知するためのフィールドである。逆に、第3ワードの“確認応答番号”フィールドは、現在のTCPコネクションにおいて送信側から受信側に送信される全バイトストリームのうち、受信側が現在何バイト目までを誤り無く受信したかを、受信側から送信側に

通知するためのフィールドである。これにより、例えば移動端末101から音声制御ホスト装置108に対して、音声データを正しい順序でかつ高い信頼性のもとで転送することが可能となる。

【0127】第4ワードの“フラグ列”フィールドには、TCPセグメントの種類を示す値が設定される。TCP通信においては、例えばコネクションの開始時又は終了時等において確認応答のための様々な制御データが通信されるが、それらの制御データの種類の“フラグ列”フィールドに設定される。

【0128】第4ワードの“ウィンドウ”フィールドは、受信側が現在何バイトのデータを連続して受信することが可能であることを示すウィンドウデータを、受信側から送信側に通知するためのフィールドである。これにより、受信側から送信側に対するデータのフロー制御が可能となり、例えば音声制御ホスト装置108の負荷が高いような場合には移動端末101に対して音声データの送信を抑制させる、というようなきめの細かい制御が可能となる。

【0129】第4ワードの“予約済”フィールドは、予約用のフィールドである。第5ワードの“チェックサム”フィールドには、TCPヘッダ及び“データ”フィールドに格納されているデータの誤りを検出するためのチェックサムデータが格納される。これにより、例えば音声制御ホスト装置108は、移動端末101から音声データを正確に受信することができる。第5ワードの“緊急ポインタ”は、緊急データ（インタラプトデータやアボートデータ等）を通信するための制御データであるが、これは本発明には特に関連しない。

【0130】第6ワードの“オプション”フィールドは、例えば送受信装置間で通信可能な最大セグメント長を指定するため等に使用されるが、これは本発明には特に関連しない。

【0131】第6ワードの“パディング”フィールドには、データ長を合わせるためのパディングデータが設定される。上述の構成を有するTCPセグメントの通信（終端）処理機能は、移動端末101においては通信部111内の通信制御部321（図3）において実現され、音声制御ホスト装置108においてはパケット送受信部115（図1）において実現される。なお、移動端末101においてCPU316が実行する制御プログラムが上記処理機能を実現するように構成されてもよい。

<発信処理>前述のように、移動端末101の制御部110内のCPU316（図3）は、図4のステップ404に対応する図5に示される送信処理のうち、移動端末101が現在図1の移動端末制御ホスト装置104に接続中でなくステップ502、507、又は511の判定がNOである場合には、ステップ503、508、又は512において、図3の通信部111内の通信制御部321に対して発信処理を依頼する。この依頼によって、

通信制御部321が実行する発信処理は、図8の動作フローチャートによって示される。

【0132】まず、ステップ801では、リンク確立フェーズが実行される。このフェーズでは、移動端末制御ホスト装置104のアクセス電話番号に対して自動的にダイヤルアップが行われ移動端末制御ホスト装置104が着信した後、リンクコントロールプロトコル(LCP)と呼ばれるプロトコルを使用し、通信に使用されるPPPフレーム(図6(a))の最大データ長の決定、エスケープされるべき非透過文字の決定、PPPフレームの“プロトコル”フィールド(図6(a))のデータ長を2オクテットから1オクテットに圧縮することの有無の決定、PPPフレームの固定値“11111111”を有する“アドレス”フィールド(図6(a))を省略(圧縮)することの有無の決定等に関するネゴシエーションが、移動端末制御ホスト装置104内の接続確立部113(図1)との間で実行される。この場合、移動端末101の通信部111内の通信制御部321と移動端末制御ホスト装置104内の接続確立部113との間の通信は、図6(a)に示されるフォーマットを有するPPPフレームを用いて、その“プロトコル”フィールドにLCPを特定する16進値“c021”を設定し、その“インフォメーションフィールド”に、必要な制御データを設定して、実行される。

【0133】次に、ステップ802においては、認証フェーズが実行される。このフェーズでは、PAP(Password Authentication Protocol)又はCHAP(ChallengeHandshake Authentication Protocol)と呼ばれる認証プロトコルを使用し、移動端末101を使用するユーザの認証が、移動端末制御ホスト装置104内の接続確立部113(図1)から移動端末101に対して実行される。これにより、移動端末制御ホスト装置104を運営するインターネットプロバイダは、移動端末101を使用するユーザが契約されたユーザであるか否かを決定できる。この場合、移動端末101の通信部111内の通信制御部321と移動端末制御ホスト装置104内の接続確立部113との間の通信は、図6(a)に示されるフォーマットを有するPPPフレームを用いて、その“プロトコル”フィールドにPAPを特定する16進値“c023”又はCHAPを特定する16進値“c223”を設定し、その“インフォメーションフィールド”に、必要な認証用データを設定して、実行される。

【0134】最後に、ステップ803では、ネットワークレイヤプロトコルフェーズが実行される。本実施の形態の場合、このフェーズでは、IPコントロールプロトコル(IPCP)と呼ばれるプロトコルを使用して、TCPヘッダ(図7(b)参照)の圧縮の有無が決定されると共に、移動端末制御ホスト装置104が割当てることができる空き(未使用)IPアドレスのうちの1つが移動端末101に対して割り当てられ、加えて、必要な経

路情報が移動端末101の通信部111内の通信制御部321(図3)と移動端末制御ホスト装置104内のルーティング部114(図1)に設定される。これ以後、移動端末101は、そのIPアドレスを使用することによって、インターネット105に接続される音声制御ホスト装置108、及びインターネット105上のユーザが希望する任意のリソースにアクセスすることが可能となる。この場合、移動端末101の通信部111内の通信制御部321と移動端末制御ホスト装置104内の接続確立部113との間の通信は、図6(a)に示されるフォーマットを有するPPPフレームを用いて、その“プロトコル”フィールドにIPCPを特定する16進値“8021”を設定し、その“インフォメーションフィールド”に、必要なIPアドレスのネゴシエーションのためのデータ等を設定して、実行される。

【0135】以上の一連の動作により、移動端末101は、移動端末制御ホスト装置104内のルーティング部114との間で通信用のTCP/IPパケットが格納されたPPPフレームを授受することが可能となり、移動端末101は、インターネット105上のリソースに自由にアクセスすることが可能になる。

【0136】なお、PHS通話時にも音声制御ホスト装置108等へのアクセスを可能とするために、移動端末101は、例えば2チャンネル同時通信機能を有するように構成することができる。

【0137】また、移動端末101の通信部111内の通信制御部321(図3)は、一定時間(例えば10分間)送受信データを検出しなかった場合に、移動端末制御ホスト装置104との間のPPPリンクを自動的に切断するように構成することができる。

<文音声認識処理に関する移動端末101の送受信処理の詳細動作>ユーザが移動端末101のタッチパネルを操作してリアルタイム又は非リアルタイムによる文音声認識処理の開始を指示した場合及びそれ以後に移動端末101が実行する送受信処理の詳細な動作について、説明する。

【0138】上述のタッチパネルの操作は、図3のタッチパネル制御部315において検出された後、制御部110内のCPU316(図3)によって、それが実行される前述した図4の動作フローチャートに対応する制御動作において、ステップ401の判定がYES、ステップ405及び406の判定がNOとなって、ステップ409の他キー入力処理が実行されることにより、検出される。更に、ステップ404の送信処理において、前述した図5のステップ501の判定がYESとなり、必要に応じてステップ503で発信処理が実行された後、ステップ504において、移動端末101の“端末識別コード”と上述の文音声認識処理の開始指示を示すキー入力処理に対応するコマンドの送信指示が、図3の通信部111内の通信制御部321に対して依頼される。



【0139】この結果、通信制御部321は、まず、図6(c)に示されるフォーマットを有するTCPセグメントを生成する。この場合に、図6(c)及び図7(b)に示されるフォーマットを有するTCPヘッダにおいて、“送信元ポート番号”フィールド及び“宛先ポート番号”フィールドには、文音声認識処理のための通信プロトコルを特定する16ビットの整数値が設定される。そして、TCPセグメントの“データ”フィールドには、移動端末101を特定する“端末識別コード”（例えばそのPHS電話番号）と、ユーザの指定に基づく文音声認識処理のリアルタイム開始要求コマンド又は文音声認識処理の非リアルタイム開始要求コマンドとが格納される。

【0140】次に、通信制御部321は、上述のTCPセグメントが“データ”フィールドに格納された図6(b)に示されるフォーマットを有するIPデータグラムを生成する。この場合に、図6(b)及び図7(a)に示されるフォーマットを有するIPヘッダにおいて、“プロトコル”フィールドには、その“データ”フィールドに格納されるTCPセグメントデータのフォーマットを規定する整数値6が設定される。また、“送信元IPアドレス”フィールドには、既に実行されている発信処理（図8のステップ803の説明を参照）によって移動端末制御ホスト装置104内の接続確立部113から移動端末101の通信部111内の通信制御部321に対して付与されたIPアドレスが設定される。更に、“宛先IPアドレス”フィールドには、音声制御ホスト装置108に割り当てられているIPアドレスが設定される。

【0141】そして、通信制御部321は、上述のIPデータグラムが“インフォメーション”フィールドに格納され、その“インフォメーション”フィールドにIPデータグラムが格納されていることを示す16進値“0021”が“プロトコル”フィールドに格納された図6(a)に示されるフォーマットを有するPPPフレームを生成し、通信制御部321内に設定されている経路情報（図8のステップ803の説明を参照）に従って、上記PPPフレームを移動端末制御ホスト装置104に送信する。以降、上述のTCPセグメント、IPデータグラム、及びPPPフレームとからなるデータ単位がインターネット105内を転送される場合に、そのデータ単位を単にTCP/IPパケットと呼ぶ。

【0142】このTCP/IPパケットは、それを構成するIPデータグラムのIPヘッダに格納されている“宛先IPアドレス”に基づいて、移動端末制御ホスト装置104内のルーティング部114とインターネット105内の特に図示しない中継ホスト装置によって、音声サービスプロバイダ内のルータ装置106まで転送された後、更に、LAN107を介して音声制御ホスト装置108内のパケット送受信部115まで転送される。

【0143】パケット送受信部115は、転送されてきたTCP/IPパケットを構成するIPデータグラムのIPヘッダの“宛先IPアドレス”フィールドに自分である音声制御ホスト装置108のIPアドレスが設定されていることを識別することによって、そのTCP/IPパケットを受信する。

【0144】そして、パケット送受信部115は、受信したTCP/IPパケットを構成するTCPセグメントの“宛先ポート番号”フィールド及び“送信元ポート番号”フィールドに文音声認識処理のための通信プロトコルを特定する16ビットの整数値が設定されていることを確認することによって、移動端末通信制御部116（図1）に対して受信通知を通知する。

【0145】この通知と共に、パケット送受信部115は、受信したTCP/IPパケットを構成するIPデータグラムのIPヘッダから“送信元IPアドレス”を取り出し、上記TCP/IPパケットを構成するTCPセグメントの“データ”フィールドから“端末識別コード”と文音声認識処理のリアルタイム開始要求コマンド又は文音声認識処理の非リアルタイム開始要求コマンドとを取り出して、それらのデータを移動端末通信制御部116に引き渡す。

【0146】この結果、後述するようにして音声制御ホスト装置108から移動端末101に対して、送信許可データが格納されたTCP/IPパケットが返信される。このTCP/IPパケットは、それを構成するIPデータグラムのIPヘッダに格納されている“宛先IPアドレス”に基づいて、音声サービスプロバイダ内のルータ装置106と、インターネット105内の特に図示しない中継ホスト装置によって、移動端末制御ホスト装置104内のルーティング部114まで転送された後、更に、PHS網103（図1）を介して移動端末101の通信部111内の通信制御部321（図3）まで転送される。

【0147】移動端末101の通信部111内の通信制御部321は、転送されてきたTCP/IPパケットを構成するIPデータグラムのIPヘッダの“宛先IPアドレス”フィールドに自分である移動端末101（に一時的又は動的）に割り当てられているIPアドレスが設定されていることを識別することによって、そのTCP/IPパケットを受信する。

【0148】そして、通信制御部321は、受信したTCP/IPパケットを構成するTCPセグメントの“宛先ポート番号”フィールド及び“送信元ポート番号”フィールドに文音声認識処理のための通信プロトコルを特定する16ビットの整数値が設定されていることを確認することにより、移動端末101の制御部110内のCPU316に対して受信通知を通知する。

【0149】この通知と共に、通信制御部321は、受信したTCP/IPパケットを構成するTCPセグメン



トの“データ”フィールドから送信許可データを取り出し、それをCPU316に引き渡す。

【0150】CPU316は、上述の受信通知と送信許可データを、前述した図4のステップ403で処理し、その送信許可データをRAM317に記憶する。移動端末101では、ユーザがタッチパネルを操作してリアルタイム又は非リアルタイムによる文音声認識処理の開始を指示することによって、CPU316が、前述した図4のステップ408で、図3の入力部109内のマイク制御部303に対して、PHS通話処理の開始指示、又は文音声認識処理を実行するためのオフライン状態での音声入力処理の開始を指示する。これにより、ユーザは、通話動作又はオフライン状態での音声入力動作によってマイク301（図2の201）からの音声の入力を開始している。

【0151】これ以後、CPU316により前述した図4のステップ401→402→403→404→401の繰返しループの1処理として実行されるステップ404の送信処理において、図5のステップ505、506の判定がYESとなり、必要に応じてステップ508で再度の発信処理が実行された後、ステップ509で、図3に示される入力部109内のマイク制御部303から制御部110内のRAM317に転送されてきている音声データの送信指示が、通信部111内の通信制御部321に対して依頼される。

【0152】この結果、通信制御部321は、まず、図6(c)に示されるフォーマットを有するTCPセグメントを生成する。この場合に、図6(c)及び図7(b)に示されるフォーマットを有するTCPヘッダにおいて、

“送信元ポート番号”フィールド及び“宛先ポート番号”フィールドには、文音声認識処理のための通信プロトコルを特定する16ビットの整数値が設定される。そして、TCPセグメントの“データ”フィールドには、図3に示される入力部109内のマイク制御部303から制御部110内のRAM317に転送されてきている音声データが格納される。

【0153】次に、通信制御部321は、上述のTCPセグメントが“データ”フィールドに格納された図6(b)に示されるフォーマットを有するIPデータグラムを生成する。この場合に、図6(b)及び図7(a)に示されるフォーマットを有するIPヘッダにおいて、“プロトコル”フォーマットには、その“データ”フィールドに格納されるTCPセグメントデータのフォーマットを規定する整数値6が設定される。また、“送信元IPアドレス”フィールドには、既に実行されている発信処理（図8のステップ803の説明を参照）によって移動端末制御ホスト装置104内の接続確立部113から移動端末101の通信部111内の通信制御部321に対して付与されたIPアドレスが設定される。更に、“宛先IPアドレス”フィールドには、音声制御ホスト装置1

08に割り当てられているIPアドレスが設定される。

【0154】そして、通信制御部321は、上述のIPデータグラムが“インフォメーション”フィールドに格納され、その“インフォメーション”フィールドにIPデータグラムが格納されていることを示す16進値“0021”が“プロトコル”フィールドに格納された図6(a)に示されるフォーマットを有するPPPフレームを生成し、通信制御部321内に設定されている経路情報（図8のステップ803の説明を参照）に従って、上記PPPフレームを移動端末制御ホスト装置104に送信する。

【0155】このTCP/IPパケットは、それを構成するIPデータグラムのIPヘッダに格納されている“宛先IPアドレス”に基づいて、移動端末制御ホスト装置104内のルーティング部114とインターネット105内の特には図示しない中継ホスト装置によって、音声サービスプロバイダ内のルータ装置106まで転送された後、更に、LAN107を介して音声制御ホスト装置108内のパケット送受信部115まで転送される。

【0156】パケット送受信部115は、転送されてきたTCP/IPパケットを構成するIPデータグラムのIPヘッダの“宛先IPアドレス”フィールドに自分である音声制御ホスト装置108のIPアドレスが設定されていることを識別することによって、そのTCP/IPパケットを受信する。

【0157】そして、パケット送受信部115は、受信したTCP/IPパケットを構成するTCPセグメントの“宛先ポート番号”フィールド及び“送信元ポート番号”フィールド文音声認識処理のための通信プロトコルを特定する16ビットの整数値が設定されていることを確認することにより、移動端末通信制御部116（図1）に対して受信通知を通知する。

【0158】この通知と共に、パケット送受信部115は、受信したTCP/IPパケットを構成するIPデータグラムのIPヘッダから“送信元IPアドレス”を取り出し、上記TCP/IPパケットを構成するTCPセグメントの“データ”フィールドから音声データを取り出して、それらのデータを移動端末通信制御部116に引き渡す。

【0159】この結果、移動端末通信制御部116は、後述するようにして文音声認識処理の制御を実行し、文音声認識部117に対して受信した音声データの認識処理を実行させる。そして、上述の音声データについて移動端末101から文音声認識処理のリアルタイム開始要求コマンドが指定されている場合には、移動端末通信制御部116は、後述するようにして、文音声認識部117からその結果として得た認識音声文章データが格納されたTCP/IPパケットを、移動端末101に対して返信する。

【0160】このTCP/IPパケットは、それを構成するIPデータグラムのIPヘッダに格納されている“宛先IPアドレス”に基づいて、音声サービスプロバイダ内のルータ装置106と、インターネット105内の特には図示しない中継ホスト装置によって、移動端末制御ホスト装置104内のルーティング部114まで転送された後、更に、PHS網103（図1）を介して移動端末101の通信部111内の通信制御部321（図3）まで転送される。

【0161】移動端末101の通信部111内の通信制御部321は、転送されてきたTCP/IPパケットを構成するIPデータグラムのIPヘッダの“宛先IPアドレス”フィールドに自分である移動端末101（一時的又は動的）に割当てられているIPアドレスが設定されていることを識別することによって、そのTCP/IPパケットを受信する。

【0162】そして、通信制御部321は、受信したTCP/IPパケットを構成するTCPセグメントの“宛先ポート番号”フィールド及び“送信元ポート番号”フィールドに文音声認識処理のための通信プロトコルを特定する16ビットの整数値が設定されていることを確認することにより、移動端末101の制御部110内のCPU316に対して受信通知を通知する。

【0163】この通知と共に、通信制御部321は、受信したTCP/IPパケットを構成するTCPセグメントの“データ”フィールドから認識音声文章データを取り出し、それをCPU316に引き渡す。

【0164】CPU316は、上述の受信通知と認識音声文章データを、前述した図4のステップ402で処理し、その認識音声文章データをLCD表示部311（図2の203）に表示する。

【0165】ユーザは、移動端末101のタッチパネルを操作することにより、音声制御ホスト装置108に対して文音声認識処理の終了を示すための、文音声認識処理の終了要求コマンドを指示することができる。

【0166】この場合に、上述のタッチパネルの操作は、図3のタッチパネル制御部315において検出された後、制御部110内のCPU316（図3）によって、それが実行される前述した図4の動作フローチャートに対応する制御動作において、ステップ401の判定がYES、ステップ405及び406の判定がNOとなつて、ステップ409の他キー入力処理が実行されることにより、検出される。更に、ステップ404の送信処理において、前述した図5のステップ501の判定がYESとなり、必要に応じてステップ503で発信処理が実行された後、ステップ504において、移動端末101の“端末識別コード”と上述の文音声認識処理の終了要求コマンドの送信指示が、図3の通信部111内の通信制御部321に対して依頼される。

【0167】この結果、通信制御部321は、まず、

“データ”フィールドに移動端末101を特定する“端末識別コード”と文音声認識処理の終了要求コマンドとが格納された図6(c)に示されるフォーマットを有するTCPセグメントを生成し、次に、そのTCPセグメントが“データ”フィールドに格納された図6(b)に示されるフォーマットを有するIPデータグラムを生成し、更に、そのIPデータグラムが“インフォメーション”フィールドに格納された図6(a)に示されるフォーマットを有するPPPフレームを生成し、それらからなるTCP/IPパケットを送信する。この場合に、TCPヘッダ（図6(c)、図7(b)）、IPヘッダ（図6(b)、図7(a)）、及び“プロトコル”フィールド（図6(a)）に設定される各情報は、前述の文音声認識処理のリアルタイム開始要求コマンド又は文音声認識処理の非リアルタイム開始要求コマンドが送信される場合に設定される各情報と同一である。

【0168】この結果、上述のTCP/IPパケットは、前述の文音声認識処理のリアルタイム開始要求コマンド等が格納されたTCP/IPパケットの場合と全く同様にして、インターネット105を介して音声制御ホスト装置108内のパケット送受信部115まで転送される。

【0169】パケット送受信部115は、前述の文音声認識処理のリアルタイム開始要求コマンド等が格納されたTCP/IPパケットが転送されてきた場合と全く同様にして、転送されてきたTCP/IPパケットを受信し、移動端末通信制御部116（図1）に対して受信通知を通知する。

【0170】この通知と共に、パケット送受信部115は、受信したTCP/IPパケットを構成するTCPセグメントの“データ”フィールドから“端末識別コード”と文音声認識処理の終了要求コマンドとを取り出して、それらのデータを移動端末通信制御部116に引き渡す。

【0171】この結果、移動端末通信制御部116は、後述するようにしてその移動端末101に対する文音声認識処理を終了する。ユーザは、文音声認識処理の非リアルタイム開始要求コマンドを指定することにより、音声制御ホスト装置108に対して、認識音声文章データを即座に移動端末101に返信することはさせずに、それを順次蓄積させることができる。この場合、ユーザは、後に、移動端末101のタッチパネルを操作することにより、音声制御ホスト装置108に蓄積されている認識音声文章データを一括して移動端末101に返信させるための、文音声認識結果の一括転送要求コマンドを指示することができる。

【0172】この場合に、上述のタッチパネルの操作は、図3のタッチパネル制御部315において検出された後、制御部110内のCPU316（図3）によって、それが実行される前述した図4の動作フローチャー

トに対応する制御動作において、ステップ401の判定がYES、ステップ405及び406の判定がNOとなつて、ステップ409の他キー入力処理が実行されることにより、検出される。更に、ステップ404の送信処理において、前述した図5のステップ501の判定がYESとなり、必要に応じてステップ503で発信処理が実行された後、ステップ504において、移動端末101の“端末識別コード”と上述の文音声認識結果の一括転送要求コマンドの送信指示が、図3の通信部111内の通信制御部321に対して依頼される。

【0173】この結果、通信制御部321はまず、“データ”フィールドに移動端末101を特定する“端末識別コード”と文音声認識結果の一括転送要求コマンドとが格納された図6(c)に示されるフォーマットを有するTCPセグメントを生成し、次に、そのTCPセグメントが“データ”フィールドに格納された図6(b)に示されるフォーマットを有するIPデータグラムを生成し、更に、そのIPデータグラムが“インフォメーション”フィールドに格納された図6(a)に示されるフォーマットを有するPPPフレームを生成し、それらからなるTCP/IPパケットを送信する。この場合に、TCPヘッダ(図6(c)、図7(b))、IPヘッダ(図6(b)、図7(a))、及び“プロトコル”フィールド(図6(a))に設定される各情報は、前述の文音声認識処理のリアルタイム開始要求コマンド又は文音声認識処理の非リアルタイム開始要求コマンドが送信される場合に設定される各情報と同一である。

【0174】この結果、上述のTCP/IPパケットは、前述の文音声認識処理のリアルタイム開始要求コマンド等が格納されたTCP/IPパケットの場合と全く同様にして、インターネット105を介して音声制御ホスト装置108内のパケット送受信部115まで転送される。

【0175】パケット送受信部115は、前述の文音声認識処理のリアルタイム開始要求コマンド等が格納されたTCP/IPパケットが転送されてきた場合と全く同様にして、転送されてきたTCP/IPパケットを受信し、移動端末通信制御部116(図1)に対して受信通知を通知する。

【0176】この通知と共に、パケット送受信部115は、受信したTCP/IPパケットを構成するIPデータグラムのIPヘッダから“送信元IPアドレス”を取り出し、上記TCP/IPパケットを構成するTCPセグメントの“データ”フィールドから“端末識別コード”と文音声認識結果の一括転送要求コマンドとを取り出して、それらのデータを移動端末通信制御部116に引き渡す。

【0177】この結果、移動端末通信制御部116は、後述するようにして、認識音声文章データが格納されたTCP/IPパケットを、移動端末101に対して一括

して転送する。

【0178】このTCP/IPパケットは、前述したようにして通信制御部321(図3)まで転送された後、通信制御部321からCPU316に、受信通知と認識音声文章データが引き渡される。

【0179】CPU316は、上述の受信通知と認識音声文章データを、前述した図4のステップ402で処理し、その認識音声文章データをLCD表示部311(図2の203)に表示する。

10 <移動端末通信制御部116と文音声認識部117の概略動作>次に、音声制御ホスト装置108内の移動端末通信制御部116と文音声認識部117の概略動作について説明する。

【0180】移動端末通信制御部116は、文音声認識処理のリアルタイム開始要求コマンド又は文音声認識処理の非リアルタイム開始要求コマンドを送信した移動端末101に割当てられている“端末識別コード”(上記コマンドを転送してきたTCPセグメントに格納されている)毎に、図13に示されるデータ構造を有する処理  
20 端末登録テーブルにエントリを登録すると共に、音声データの受信用のバッファファイル(音声バッファファイル)と認識音声文章データの送信用のバッファファイル(文章バッファファイル)とを音声制御ホスト装置108が管理するファイルシステム上に作成する。また、移動端末通信制御部116は、上記エントリとファイルの登録に成功すると、上記コマンドを転送してきたIPデータグラムに格納されていた“送信元IPアドレス”の移動端末101に向けて、送信許可データを返信する。

【0181】移動端末通信制御部116は、それ以後移動端末101から受信した音声データを、その“送信元IPアドレス”(それを転送してきたIPデータグラムに格納されている)に対応する処理端末登録テーブルのエントリから特定される音声バッファファイルに追加書き込みする。

【0182】文音声認識部117は、図13に示される処理端末登録テーブルのエントリ毎に、各エントリから特定される音声バッファファイルに音声データが受信されていればそれに対して文音声認識処理を実行し、その結果得られる認識音声文章データを上記各エントリに対応する文章バッファファイルに追加書き込みする。

【0183】移動端末通信制御部116は、リアルタイム要求が設定されている処理端末登録テーブルのエントリ毎に、各エントリから特定される文章バッファファイルに認識音声文章データが得られていれば、それを各エントリに登録されている“送信元IPアドレス”の移動  
40 端末101に向けて返信する。

【0184】移動端末通信制御部116は、文音声認識処理の終了要求コマンドを受信しかつリアルタイム要求が設定されている処理端末登録テーブルのエントリ、又はリアルタイム要求が設定されておりかつ最終アクセス  
50

時刻が現在時刻から一定時間前の時刻よりも更に前の時刻である処理端末登録テーブルのエントリについて、そのエントリの内容を削除し、それから特定される音声／文章バッファファイルを削除する。

【0185】移動端末通信制御部116は、リアルタイム要求が設定されていない（即ち、非リアルタイム要求が設定されている）処理端末登録テーブルのエントリについて、そのエントリの内容とそれから特定される音声／文章バッファファイルは、対応する移動端末101との通信終了後も保存し、後に、その移動端末101から文音声認識結果の一括転送要求コマンドを受信した時点で、その移動端末101の“端末識別コード”に対応する処理端末登録テーブルのエントリから特定される文章バッファファイルに保存されていた認識音声文章データを、上記コマンドを転送してきたIPデータグラムに格納されていた“送信元IPアドレス”の移動端末101に向けて送信する。

＜移動端末通信制御部116の詳細動作＞図9～図12は、上記機能を実現するために、移動端末通信制御部116が実行する制御動作を示す動作フローチャートである。この動作フローチャートは、移動端末通信制御部116を制御する特には図示しないプロセッサが、特には図示しない制御プログラムを実行する動作として実現される。

【0186】まず、ステップ901で、音声制御ホスト装置108内のパケット送受信部115（図1）から受信通知が通知されたか否かが判定される。前述したように、パケット送受信部115は、インターネット105から転送されてきたTCP/IPパケットを構成するIPデータグラムのIPヘッダの“宛先IPアドレス”フィールドに自分である音声制御ホスト装置108のIPアドレスが設定されていることを識別することにより、そのTCP/IPパケットを受信し、かつ、それを構成するTCPセグメントの“宛先ポート番号”フィールド及び“送信元ポート番号”フィールドに文音声認識処理のための通信プロトコルを特定する16ビットの整数値が設定されていることを確認することによって、移動端末通信制御部116に対して受信通知を通知する。この受信通知は、文音声認識処理のリアルタイム開始要求コマンド、文音声認識処理の非リアルタイム開始要求コマンド、文音声認識処理の対象である音声データ、文音声認識処理の終了要求コマンド、又は文音声認識結果の一括転送要求コマンドの何れかに関する受信通知である。パケット送受信部115から受信通知が通知されステップ901の判定がYESとなると、ステップ902で、パケット送受信部115から受信通知と共に引き渡されたデータが取り込まれる。この場合に、受信通知が、文音声認識処理のリアルタイム開始要求コマンド、文音声認識処理の非リアルタイム開始要求コマンド、又は文音声認識結果の一括転送要求コマンドの何れかの受信通知

である場合には、“送信元IPアドレス”と“端末識別コード”と上記コマンドとが取り込まれる。また、受信通知が、音声データの受信通知である場合には、“送信元IPアドレス”と音声データとが取り込まれる。更に、受信通知が、文音声認識処理の終了要求コマンドの受信通知である場合には、“端末識別コード”とそのコマンドとが取り込まれる。

【0187】ステップ902の処理の後に、図9のステップ903、図10のステップ909、図10のステップ911、又は図11のステップ915の判定が順に検査され、何れかの判定結果がYESとなる。即ち、ステップ902でパケット送受信部115から引き渡されたデータが、文音声認識処理のリアルタイム開始要求コマンド又は文音声認識処理の非リアルタイム開始要求コマンドに関するものである場合はステップ903の判定がYESとなってステップ904～908が実行され、音声データに関するものである場合は図10のステップ909の判定がYESとなってステップ910が実行され、文音声認識処理の終了要求コマンドに関するものである場合には図10のステップ911の判定がYESとなってステップ912～914が実行され、文音声認識結果の一括転送要求コマンドに関するものである場合には図11のステップ915の判定がYESとなってステップ916～918が実行される。

【0188】パケット送受信部115から受信通知が通知されておらずステップ901の判定がNOの場合、又は上述の各コマンド又は音声データの受信に対応する処理の後は、図12のステップ919と920で認識音声文章データの送信処理が実行され、それに続くステップ921及び922で最終アクセス時刻が一定時間以上前である移動端末101との通信を終了させるための処理が行われた後、再び図9のステップ901の判定処理に戻る。

【0189】ステップ901の判定がYESであり、ステップ902でパケット送受信部115から引き渡されたデータが文音声認識処理のリアルタイム開始要求コマンド又は文音声認識処理の非リアルタイム開始要求コマンドに関するものである場合において、ステップ903の判定がYESとなって実行されるステップ904～908の処理について説明する。

【0190】まず、ステップ904では、ステップ902でパケット送受信部115から引き渡された“端末識別コード”に対応するエントリが、処理端末登録テーブルに登録されているか否かが判定される。

【0191】移動端末101から初めて文音声認識処理のリアルタイム開始要求コマンド又は文音声認識処理の非リアルタイム開始要求コマンドが指定された場合には、この判定はNOとなる。

【0192】その結果、ステップ905では、音声データの受信用のバッファファイルである音声バッファファ

10

20

30

40

50

イルと、認識音声文章データの送信用のバッファファイルである文章バッファファイルとが、音声制御ホスト装置108が管理するファイルシステム上に作成される。

【0193】次に、ステップ906では、移動端末通信制御部116内の特には図示しないメモリに記憶される図13に示されるデータ構造を有する処理端末登録テーブルに、1つのエントリ（横1行のデータ組）が確保される。そして、そのエントリに、“端末識別コード”と、“送信元IPアドレス”と、リアルタイム要求の有無と、最終アクセス時刻と、音声バッファファイル名と、文章バッファファイル名とが、登録される。“端末識別コード”は、ステップ902においてパケット送受信部115から引き渡されたデータであり、移動端末101から転送されてきたTCP/IPパケットを構成するTCPセグメントの“データ”フィールドに格納されていたものである（図6(c)参照）。“送信元IPアドレス”は、やはりステップ902においてパケット送受信部115から引き渡されたデータであり、移動端末101から転送されてきたTCP/IPパケットを構成するIPデータグラムのIPヘッダに格納されていたものである（図6(b)、図7(a)参照）。リアルタイム要求の有無は、ステップ902においてパケット送受信部115から引き渡されたコマンドが、文音声認識処理のリアルタイム開始要求コマンドである場合には“有り”（値“1”）に設定され、文音声認識処理の非リアルタイム開始要求コマンドである場合には“無し”（値“0”）に設定される。最終アクセス時刻には、現在時刻が設定される。音声バッファファイル名と文章バッファファイル名は、ステップ905で作成された各ファイルを示すファイル名である。

【0194】次に、前述したように、移動端末101から初めて文音声認識処理のリアルタイム開始要求コマンド又は文音声認識処理の非リアルタイム開始要求コマンドが指定された場合には、処理端末登録テーブルに対応するエントリは存在しないため、前述したステップ904の判定はYESとなる。また、前述したように、移動端末101から文音声認識処理のリアルタイム開始要求コマンドが指定されて文音声認識処理の実行が開始された後に、その移動端末101から文音声認識処理の終了要求コマンドを受信した場合又は最終アクセス時刻が現在時刻から一定時間前の時刻よりも更に前の時刻である（即ち一定時間アクセスがない）場合には、処理端末登録テーブル上の対応するエントリは削除される。

【0195】しかし、移動端末101から文音声認識処理のリアルタイム開始要求コマンドが指定されて文音声認識処理の実行が開始された後に、処理端末登録テーブル上の対応するエントリが削除されないうちに、再び同じ移動端末101から文音声認識処理のリアルタイム開始要求コマンド又は文音声認識処理の非リアルタイム開始要求コマンドが指定された場合は、ステップ904の

判定はYESとなる。また、前回、移動端末101から文音声認識処理の非リアルタイム開始要求コマンドが指定された文音声認識処理が実行された後、まだ、その移動端末101から文音声認識結果の一括転送要求コマンドが指定されていないうちに、再び同じ移動端末101から文音声認識処理のリアルタイム開始要求コマンド又は文音声認識処理の非リアルタイム開始要求コマンドが指定された場合にも、ステップ904の判定はYESとなる（後述する図11のステップ917参照）。

【0196】このような場合には、処理端末登録テーブル上の削除されていない前回と同じエントリが使用される。そして、ステップ907において、処理端末登録テーブル上の上記エントリに記憶されている“送信元IPアドレス”とリアルタイム要求の有無のみが、図9のステップ902においてパケット送受信部115から引き渡された新しいデータに更新され、また、最終アクセス時刻が現在時刻に更新される。

【0197】前述したステップ906又は上記ステップ907の処理の後、ステップ908では、ステップ902でパケット送受信部115から引き渡され処理端末登録テーブルの上記エントリに登録された“送信元IPアドレス”に向けて、送信許可データが返信される。

【0198】具体的には、移動端末通信制御部116は、“送信元IPアドレス”への送信許可データの返信を、パケット送受信部115（図1）に対して依頼する。この結果、パケット送受信部115は、まず、図6(c)に示されるフォーマットを有するTCPセグメントを生成する。この場合、図6(c)及び図7(b)に示されるフォーマットを有するTCPヘッダにおいて、“送信元ポート番号”フィールド及び“宛先ポート番号”フィールドには、文音声認識処理のための通信プロトコルを特定する16ビットの整数値が設定される。そして、TCPセグメントの“データ”フィールドには、送信許可データが格納される。

【0199】次に、パケット送受信部115は、上述のTCPセグメントが“データ”フィールドに格納された図6(b)に示されるフォーマットを有するIPデータグラムを生成する。この場合に、図6(b)及び図7(a)に示されるフォーマットを有するIPヘッダにおいて、

“プロトコル”フォーマットには、その“データ”フィールドに格納されるTCPセグメントデータのフォーマットを規定する整数値6が設定される。また、“送信元IPアドレス”フィールドには、音声制御ホスト装置108に割当てられているIPアドレスが設定される。更に、“宛先IPアドレス”フィールドには、図9のステップ902でパケット送受信部115から引き渡された“送信元IPアドレス”が設定される。

【0200】そして、パケット送受信部115は、上述のIPデータグラムが格納されたLAN107上のプロトコルに従ったフレームを生成し、それをLAN107

に送出する。例えば、LAN107がイーサネット方式によるローカルエリアネットワークであれば、上記フレームは、イーサネットフレームである。

【0201】上記フレームとIPデータグラムとTCPセグメントとから構成されるTCP/IPパケットは、それを構成するIPデータグラムのIPヘッダに格納されている“宛先IPアドレス”に基づいて、ルータ装置106及びインターネット105を介して移動端末制御ホスト装置104まで転送された後、更に、PHS網103及び無線基地（又は有線接続装置）102を介して、移動端末101の通信部111内の通信制御部321（図3）まで転送される。

【0202】これ以降、移動端末101から音声制御ホスト装置108へは、前述したようにして、音声データが転送されてくる。ステップ908の処理の後、図12のステップ919と920で認識音声文章データの送信処理が実行され、それに続くステップ921及び922で最終アクセス時刻が一定時間以上前である移動端末101との通信を終了させるための処理が行われた後、再び図9のステップ901の判定処理に戻る。

【0203】次に、図9のステップ901の判定がYESであり、ステップ902でパケット送受信部115から引き渡されたデータが音声データである場合において、図10のステップ909の判定がYESとなって実行されるステップ910の処理について説明する。

【0204】即ち、ステップ910では、図9のステップ902でパケット送受信部115から引き渡されたのと同じ“送信元IPアドレス”が記憶されている処理端末登録テーブル（図13）のエントリが検索され、該当するエントリに記憶されている音声バッファファイル名に対応する音声バッファファイル（図9のステップ906参照）に、図9のステップ902でパケット送受信部115から引き渡された音声データが追加書き込みされる。なお、追加書き込み時の音声バッファファイルのサイズは、音声制御ホスト装置108が管理するファイルシステムによって自動的に調整される。

【0205】また、ステップ910では、上記該当するエントリに記憶されている最終アクセス時刻が、現在時刻に更新される。このようにして、移動端末101毎（“端末識別コード”毎）の音声バッファファイルを介して、移動端末通信制御部116から文音声認識部117（図1）に音声データが引き渡される。即ち、文音声認識部117は、後述するように、図13に示される処理端末登録テーブルのエントリ毎に、各エントリから特定される音声バッファファイルに音声データが受信されればそれに対して文音声認識処理を実行し、その結果得られる認識音声文章データを上記各エントリに対応する文章バッファファイルに追加書き込みすることになる。

【0206】ステップ910の処理の後、図12の

ステップ919と920で認識音声文章データの送信処理が実行され、それに続くステップ921及び922で最終アクセス時刻が一定時間以上前である移動端末101との通信を終了させるための処理が行われた後、再び図9のステップ901の判定処理に戻る。

【0207】次に、図9のステップ901の判定がYESであり、ステップ902でパケット送受信部115から引き渡されたデータが文音声認識処理の終了要求コマンドに関するものである場合において、図10のステップ911の判定がYESとなって実行されるステップ912～914の処理について説明する。

【0208】まず、ステップ912では、図9のステップ902でパケット送受信部115から引き渡されたのと同じ“端末識別コード”が記憶されている処理端末登録テーブル（図13）のエントリに、リアルタイム要求として“有り”（値“1”）が記憶されているか否かが判定される。

【0209】上記エントリにリアルタイム要求として“有り”（値“1”）が記憶されていない、即ちリアルタイム要求として“無し”（値“0”）が記憶されており、移動端末101から最初に文音声認識処理の非リアルタイム開始要求コマンドが指示されていた場合には、文音声認識処理の終了要求コマンドの受信後も処理端末登録テーブルのエントリ、文章バッファファイル、及び音声バッファファイル（この内容は通常文音声認識処理が終了した時点で空となる）の各内容は保持されるため、ステップ912の判定がNOとなってステップ913及び914の処理は実行されず、図12のステップ919と920で認識音声文章データの送信処理が実行され、それに続くステップ921及び922で最終アクセス時刻が一定時間以上前である移動端末101との通信を終了させるための処理が行われた後、再び図9のステップ901の判定処理に戻る。

【0210】一方、前記エントリにリアルタイム要求として“有り”（値“1”）が記憶されている、即ち移動端末101から最初に文音声認識処理のリアルタイム開始要求コマンドが指示されていた場合は、ステップ912の判定がYESとなる。

【0211】この結果、まず、ステップ913で、図9のステップ902でパケット送受信部115から引き渡されたのと同じ“端末識別コード”が記憶されている処理端末登録テーブル（図13）のエントリの内容が全て削除される。

【0212】また、ステップ914で、上記エントリに記憶されていた音声バッファファイル名に対応する音声バッファファイル及び文章バッファファイル名に対応する文章バッファファイルが、音声制御ホスト装置108が管理するファイルシステム上から削除される。

【0213】ステップ914の処理の後、図12のステップ919と920で認識音声文章データの送信処理



が実行され、それに続くステップ921及び922で最終アクセス時刻が一定時間以上前である移動端末101との通信を終了させるための処理が行われた後、再び図9のステップ901の判定処理に戻る。

【0214】続いて、図9のステップ901の判定がYESであり、ステップ902でパケット送受信部115から引き渡されたデータが音声認識結果の一括転送要求コマンドに関するものである場合において、図11のステップ915の判定がYESとなって実行されるステップ916～918の処理について説明する。

【0215】まず、ステップ916では、図9のステップ902でパケット送受信部115から引き渡されたのと同じ“端末識別コード”が記憶されている処理端末登録テーブル(図13)のエントリから得られる文章バッファファイル名に対応する文章バッファファイルから、認識音声文章データが一括して読み出され、それが1つ以上のTCP/IPパケットに格納され、図9のステップ902でパケット送受信部115から引き渡された“送信元IPアドレス”に向けて送信される。

【0216】具体的には、移動端末通信制御部116は、“送信元IPアドレス”への認識音声文章データの送信を、パケット送受信部115に対し依頼する。この結果、パケット送受信部115は、まず、図6(c)に示されるフォーマットを有するTCPセグメントを生成する。この場合、図6(c)及び図7(b)に示されるフォーマットを有するTCPヘッダにおいて、“送信元ポート番号”フィールド及び“宛先ポート番号”フィールドには、音声認識処理のための通信プロトコルを特定する16ビットの整数値が設定される。そして、TCPセグメントの“データ”フィールドには、認識音声文章データが格納される。

【0217】次に、パケット送受信部115は、上述のTCPセグメントが“データ”フィールドに格納された図6(b)に示されるフォーマットを有するIPデータグラムを生成する。この場合に、図6(b)及び図7(a)に示されるフォーマットを有するIPヘッダにおいて、“プロトコル”フォーマットには、その“データ”フィールドに格納されるTCPセグメントデータのフォーマットを規定する整数値6が設定される。また、“送信元IPアドレス”フィールドには、音声制御ホスト装置108に割当てられているIPアドレスが設定される。更に、“宛先IPアドレス”フィールドには、図9のステップ902でパケット送受信部115から引き渡された“送信元IPアドレス”が設定される。

【0218】そして、パケット送受信部115は、上述のIPデータグラムが格納されたLAN107上のプロトコルに従ったフレームを生成し、それをLAN107に送出する。

【0219】上記フレームとIPデータグラムとTCPセグメントとから構成されるTCP/IPパケットは、

それを構成するIPデータグラムのIPヘッダに格納されている“宛先IPアドレス”に基づいて、ルータ装置106及びインターネット105を介して移動端末制御ホスト装置104まで転送された後、更に、PHS網103及び無線基地(又は有線接続装置)102を介して、移動端末101の通信部111内の通信制御部321(図3)まで転送され、前述したように、移動端末101のLCD表示部311(図2の203)に表示される。

10 【0220】次に、ステップ917では、図9のステップ902でパケット送受信部115から引き渡されたのと同じ“端末識別コード”が記憶されている処理端末登録テーブル(図13)のエントリの内容が全て削除される。

【0221】また、ステップ918で、上記エントリに記憶されていた音声バッファファイル名に対応する音声バッファファイル及び文章バッファファイル名に対応する文章バッファファイルが、音声制御ホスト装置108が管理するファイルシステム上から削除される。

20 【0222】ステップ918の処理の後には、図12のステップ919と920で認識音声文章データの送信処理が実行され、それに続くステップ921及び922で最終アクセス時刻が一定時間以上前である移動端末101との通信を終了させるための処理が行われた後、再び図9のステップ901の判定処理に戻る。

【0223】パケット送受信部115から受信通知が通知されておらず図9のステップ901の判定がNOの場合、又は上述の各コマンド又は音声データの受信に対応する処理の後に実行される、図12のステップ919と920の処理、及びそれに続くステップ921と922の処理について説明する。

【0224】これらの処理において、音声認識部117から得られている認識音声文章データの送信処理が実行される。まず、ステップ919では、処理端末登録テーブル(図13)において、リアルタイム要求として“有り”(値“1”)が記憶されており、かつ文章バッファファイル名に対応する文章バッファファイルに認識音声文章データが存在するエントリがあるか否かが判定される。

40 【0225】そのようなエントリが無くステップ919の判定がNOの場合には、ステップ920での認識音声文章データの送信処理は実行されずに、ステップ921及び922の処理に進む。

【0226】上述のようなエントリが1つ以上存在しステップ919の判定がYESの場合には、ステップ920で、該当するエントリ毎に、そのエントリに記憶されている“送信元IPアドレス”に向けて、そのエントリに記憶されている文章バッファファイル名に対応する文章バッファファイル内の認識音声文章データが送信され、その送信された認識音声文章データが上記文章バッ



ファファイルから削除される。なお、削除時の文章バッファファイルのサイズは、音声制御ホスト装置108が管理するファイルシステムによって自動的に調整される。

【0227】上述のステップ920の処理の後又はステップ919の判定がNOである場合に、ステップ921が実行される。ここでは、処理端末登録テーブル(図13)のエントリのうち、リアルタイム要求として“有” (値“1”) が設定されており、かつ最終アクセス時刻が現在時刻から一定時間前の時刻よりも更に前の時刻であるエントリが検出され、そのエントリの内容が全て削除される。

【0228】また、ステップ922で、上記エントリに記憶されていた音声バッファファイル名に対応する音声バッファファイル及び文章バッファファイル名に対応する文章バッファファイルが、音声制御ホスト装置108が管理するファイルシステム上から削除される。

【0229】ステップ922の処理の後、再び図9のステップ901の判定処理に戻る。

<文音声認識部117の詳細動作>図14は、文音声認識部117の機能ブロック図である。

【0230】この文音声認識部117は、前述したように、図13に示される処理端末登録テーブルのエントリ毎に、各エントリから特定される音声バッファファイルに音声データが受信されていればそれに対して文音声認識処理を実行し、その結果得られる認識音声文章データを上記各エントリに対応する文章バッファファイルに追加書き込みする。

【0231】上述のエントリ毎の音声バッファファイルからの音声データの読出しと文章バッファファイルへの認識音声文章データの書き込みは、図14の入出力制御部1409が制御する。まず、この入出力制御部1409の制御動作につき説明する。図15は、入出力制御部1409が実行する制御動作を示す動作フローチャートである。この動作フローチャートは、入出力制御部1409を制御する特には図示しないプロセッサが、特には図示しない制御プログラムを実行する動作として実現される。

【0232】まず、ステップ1501では、処理端末登録テーブル(図13)において、音声バッファファイル名に対応する音声バッファファイルに音声データが記憶されているエントリが存在するかどうか判定される。

【0233】そのようなエントリが存在しステップ1501の判定がYESならば、ステップ1502で、該当するエントリ毎に、そのエントリに記憶されている“端末識別コード”と、そのエントリに記憶されている音声バッファファイル名に対応する音声バッファファイル上の音声データとが、図14の入力バッファキュー1401に書き込まれ、その音声データが音声バッファファイルから削除される。

【0234】入力バッファキュー1401は、それがキューイングしている音声データを、音声区間検出部1402に順次流し込む機能を有する。音声区間検出部1402以降に接続されている音声分析部1403、音素認識部1404、単語認識部1406、及び文章認識部1407は、データ処理パイプラインを形成しており、相互に独立して、入力データを処理する機能を有する。また、1402~1407の各部分は、現在処理している音声データに対応する“端末識別コード”(入力バッファキュー1401から入力される)を認識することができる。従って、最終的に文章認識部1407から出力バッファキュー1408へは、“端末識別コード”と認識音声文章データとの組が出力されることになる。

【0235】ステップ1502の処理の後又はステップ1501の判定がNOの場合には、ステップ1503で、図14の出力バッファキュー1408に、“端末識別コード”と認識音声文章データの組が得られているかどうか判定される。

【0236】そのような組が得られておりステップ1503の判定がYESならば、ステップ1504で、出力バッファキュー1408内の組毎に、その組の“端末識別コード”に対応する処理端末登録テーブルのエントリについて、そのエントリに記憶されている文章バッファファイル名に対応する文章バッファファイルに、出力バッファキュー1408内の組の認識音声文章データが追加書き込みされる。

【0237】ステップ1504の処理の後又はステップ1503の判定がNOの場合には、再びステップ1501の判定処理が実行される。以上のようにして文音声認識部117は、流れ作業的に効率良く、複数の移動端末101から要求された音声データに対する文音声認識処理を実行することができる。

【0238】次に、文音声認識処理を実現するための1402~1407の各部分の機能につき、以下に説明する。なお、以下に説明する各方式は、例えば、文献「電子・情報工学入門シリーズ2 音響・音声工学」(古井著、近代科学社)第14章を参照することにより、実現することができる。

【0239】音声区間検出部1402は、入力バッファキュー1401から入力される音声データのサンプル時系列について、音声が存在する区間を検出する。より具体的には、音声区間検出部1402は、所定サンプル(例えば8kHzサンプリングデータについて32乃至256サンプル)ずつの平均パワー(電力)を計算し、その平均パワーが所定の閾値を超えた状態が所定回数以上連続して続く区間を、音声区間として検出する。これにより、音声が存在しない区間で文音声認識が誤認識されてしまうのを防ぐことができる。

【0240】音声分析部1403は、音声区間検出部1402から出力される音声データについて、その特徴分

10

20

30

40

50

析を行うことによって、特徴量パラメータベクトルを検出する。音声分析方式としては、以下の周知の分析方式の何れかを採用することができる。

(1) 音声データ時系列を入力とする帯域フィルタバンクの各出力を平滑化し、それらの平滑化された各出力を特徴量パラメータベクトルの要素とする方式。

(2) 連続する所定サンプルずつの音声データ時系列を入力とする高速フーリエ変換(FFT)によって計算した各短時間スペクトル成分を平滑化し、それらの平滑化された各成分値を特徴量パラメータベクトルの要素とする方式。

(3) 連続する所定サンプルずつの音声データ時系列を入力とするケプストラム分析によってケプストラム係数群を計算し、それらを特徴量パラメータベクトルの要素とする方式。

(4) 上記(3)のケプストラム係数群に加えて、それらに対する $\Delta$ (デルタ)ケプストラム(ケプストラムの微係数)群を計算し、それらを特徴量パラメータベクトルの要素に加える方式。

(5) 連続する所定サンプルずつの音声データ時系列を入力とする線形予測分析(LPC分析、更に具体的には線スペクトル対分析:LSP分析)によって、LPC

(LSP)係数群を計算し、それらを特徴量パラメータベクトルの要素とする方式。

(6) 連続する所定サンプルずつの音声データ時系列を入力とする自己相関分析によって自己相関関数を計算し、それらに基づいて検出される音声のピッチ基本周波数パターンを特徴量パラメータベクトルの1つの要素に加える方式。

次に、音素認識部1404は、所定フレーム周期(所定サンプル)毎に音声分析部1403から出力される特徴量パラメータベクトルと、音素標準パターン辞書1405に蓄積されている各音素の特徴量パラメータベクトルの標準パターンとの類似度(距離)を計算し、その結果所定フレーム周期毎に得られる類似度の高い音素の組をその類似度と共に音素ラティスデータとして出力する。音素認識部1404は、音素の認識誤りの発生を回避するために、所定フレーム周期毎に最終的な音素を決定することはせずに、音素候補を表にした音素ラティスデータの形式で結果データを出力する。

【0241】単語認識部1406は、所定フレーム周期毎に音素認識部1404から出力される音素ラティスデータを入力として、所定フレーム周期毎に単語候補を表にして単語ラティスデータを出力する。単語認識方式としては、以下の周知の分析方式の何れかを採用することができる。

(1) 単語認識部1406は、音素認識部1404から出力される複数のフレーム周期にまたがる音素ラティスデータの時系列と、単語辞書に蓄積されている全音素標準パターン系列とで、時間正規化(DPマッチング or

DTW: Dynamic Time Warping)を実行し、単語ラティスデータを出力する。この場合も、単語認識部1406は、単語の認識誤りの発生を回避するために、所定フレーム周期毎に最終的な単語を決定することはせずに、単語候補を表にした単語ラティスデータの形式で結果データを出力する。

(2) 単語認識部1406は、HMM(Hidden Markov Model)によって、全単語をモデル化し、音素認識部1404から出力される複数のフレーム周期にまたがる音素ラティスデータの時系列をHMM分析部に入力し、生起確率の大きいものから複数個のモデルに対応する各単語を、単語候補である単語ラティスデータとして出力する。最後に、文章認識部1407は、その第1段処理として、単語認識部1406から出力される単語ラティスデータを順次入力し、日本語(英語でもよい)の文節構造に関する文節内文法(語順規則)に従って、種々の文節の可能性を文節ラティスデータとして算出する。そして、文章認識部1407は、その第2段処理として、文節間文法に従って文節間の意味的な係り受けを解析し、認識音声文章データを決定し、それを、入力バッファキュー1401から順次伝達されてきた“端末識別コード”と対について、出力バッファキュー1408に書き込む。

<他の実施の形態>以上説明した実施の形態では、移動端末101は、PHS端末であって、移動端末101と音声制御ホスト装置108とは、PHS網103とインターネット105を介して接続されている。しかし、本発明は、これに限られるものではなく、無線又は有線によって間接的又は直接的に音声制御ホスト装置108に接続される形態であれば、どのような形態であっても本発明をそれに適用することができる。

【0242】

【発明の効果】本発明によれば、移動端末は、高度な音声認識環境を設備する必要がなく実用的な精度を有する音声認識機能の提供を低コストで受けることが可能となる。

【0243】また、本発明によれば、現在全国的及び全世界的に普及しつつあるパーソナルハンディホンシステム通信網及びインターネットを経由することによって、実用的な精度を有する音声認識機能の提供をより低コスト及び手軽に受けることができると同時に、本発明が提供する機能とパーソナルハンディホンシステム通話機能及びインターネットアクセス機能とを、シームレスに結合することが可能となる。

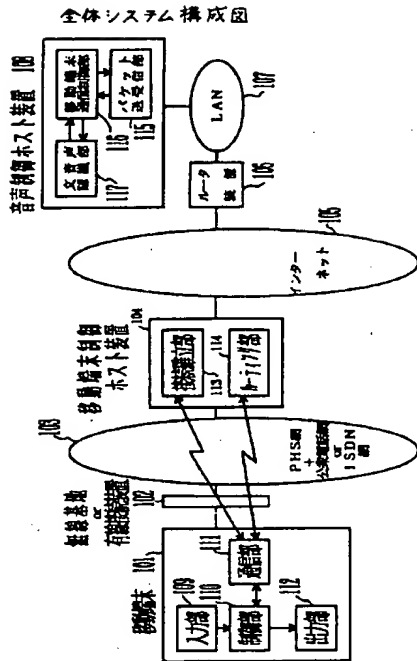
【0244】更に、本発明によれば、移動端末と音声制御ホスト装置とを全世界的に容易に特定できると共に、音声認識処理サービスと他の情報処理サービスとの共存を容易に実現することが可能となる。

【0245】加えて、本発明によれば、ホスト装置側の負荷分散を容易に実現することが可能となる。

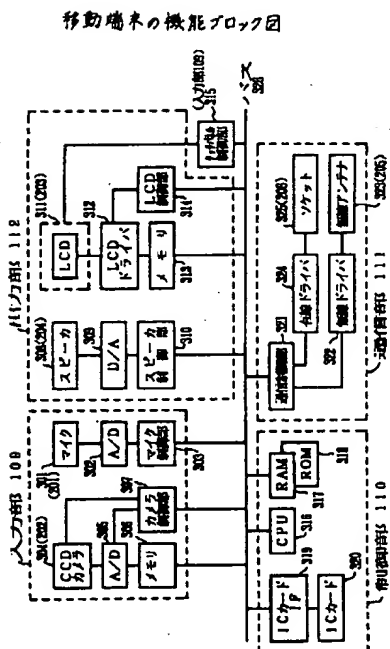
【図面の簡単な説明】  
 【図1】全システム構成図である。  
 【図2】移動端末の外観図である。  
 【図3】移動端末の機能ブロック図である。  
 【図4】移動端末の処理の全体動作フローチャートである。  
 【図5】送信処理の動作フローチャートである。  
 【図6】通信データのフォーマット図である。  
 【図7】IPヘッダとTCPヘッダのフォーマット図である。  
 【図8】PPPを用いた発信処理の動作フローチャートである。  
 【図9】移動端末通信制御部の動作フローチャート（その1）である。  
 【図10】移動端末通信制御部の動作フローチャート（その2）である。  
 【図11】移動端末通信制御部の動作フローチャート（その3）である。  
 【図12】移動端末通信制御部の動作フローチャート（その4）である。  
 【図13】処理端末登録テーブルのデータ構成図である。  
 【図14】文音声認識部の構成図である。  
 【図15】入出力制御部の動作フローチャートである。  
 【符号の説明】  
 101 移動端末  
 102 無線基地（有線接続装置）  
 103 PHS網（公衆電話網、ISDN網）  
 104 移動端末制御ホスト装置  
 105 インターネット  
 106 ルータ装置  
 107 LAN（ローカルエリアネットワーク）  
 108 音声制御ホスト装置  
 109 入力部  
 110 制御部  
 111 通信部  
 112 出力部  
 113 接続確立部

114 ルーティング部  
 115 パケット送受信部  
 116 移動端末通信制御部  
 117 文音声認識部  
 201、301 マイク  
 202、304 カメラ（CCDカメラ）  
 203、311 LCD表示部  
 204、308 スピーカ  
 205、323 無線アンテナ  
 10 206、325 ソケット（通信用）  
 207 ICカードスロット  
 208 光送受信機（光通信用）  
 302、305 A/D変換部  
 303 マイク制御部  
 306、313 メモリ  
 307 カメラ制御部  
 309 D/A変換部  
 310 スピーカ制御部  
 312 LCDドライバ  
 20 314 LCD制御部  
 315 タッチパネル制御部  
 316 CPU  
 317 RAM  
 318 ROM  
 319 ICカードインタフェース部  
 320 ICカード  
 321 通信制御部  
 322 無線ドライバ  
 324 有線ドライバ  
 30 1401 入力バッファキュー  
 1402 音声区間検出部  
 1403 音声分析部  
 1404 音素認識部  
 1405 音素標準パターン辞書  
 1406 単語認識部  
 1407 文章認識部  
 1408 出力バッファキュー  
 1409 入出力制御部

【図1】

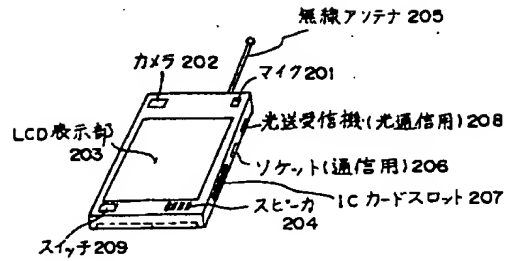


【図3】



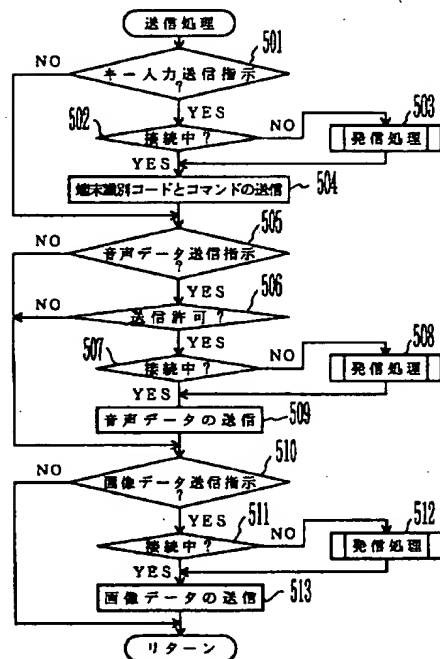
【図2】

移動端末の外観図



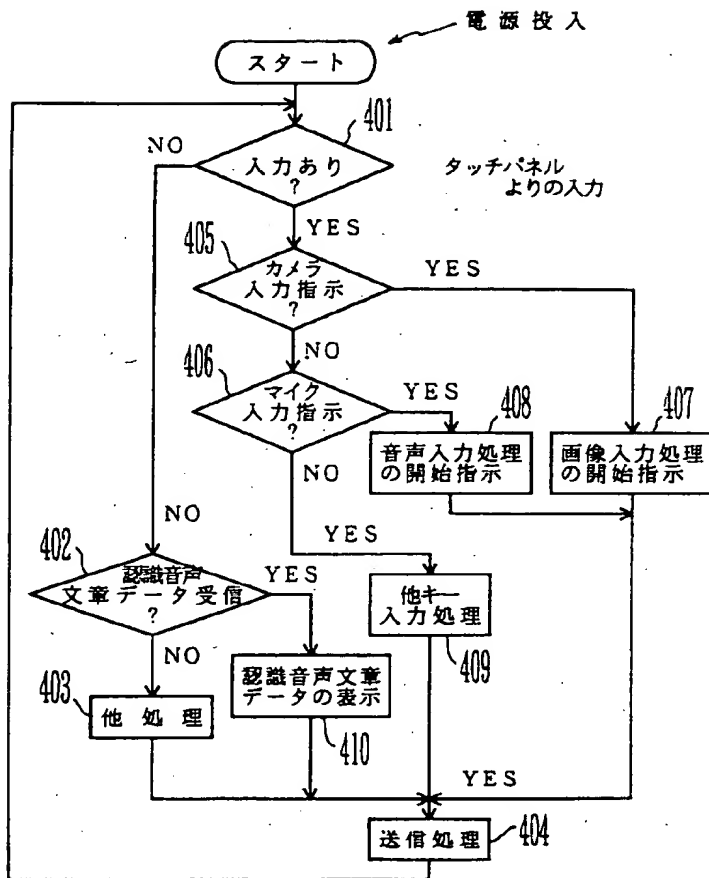
【図5】

送信処理の動作フローチャート



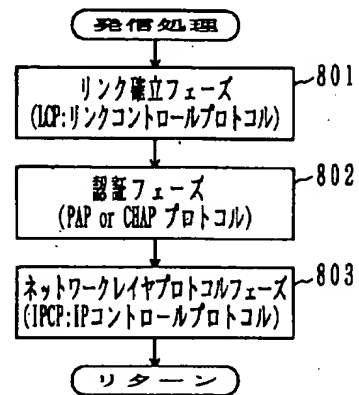
【図4】

## 移動端末の処理の全体動作フローチャート



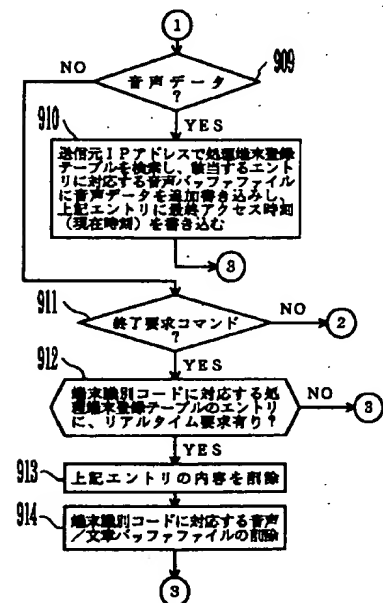
【図8】

## PPPを用いた発信処理の動作フローチャート



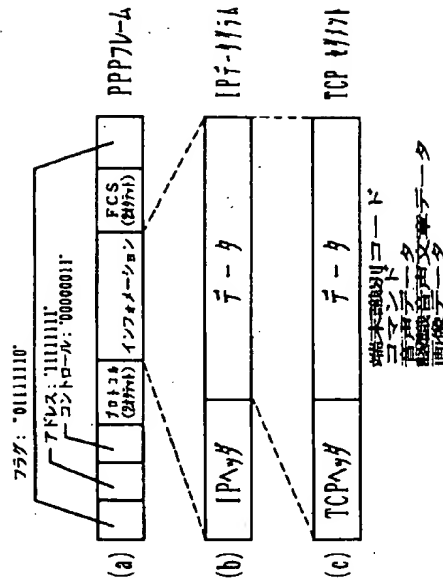
【図10】

## 移動端末通信制御部の動作フローチャート(その2)



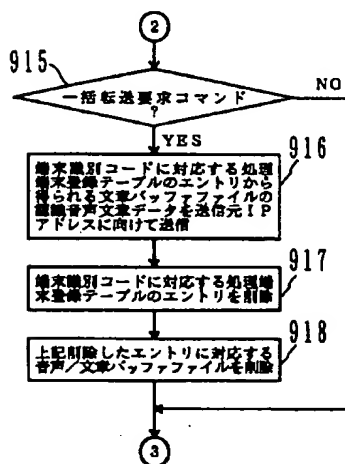
【図6】

通信データのフォーマット図



【図11】

移動端末通信制御部の動作フローチャート(その3)



【図7】

IPヘッダとTCPヘッダのフォーマット図

(a) IPヘッダ

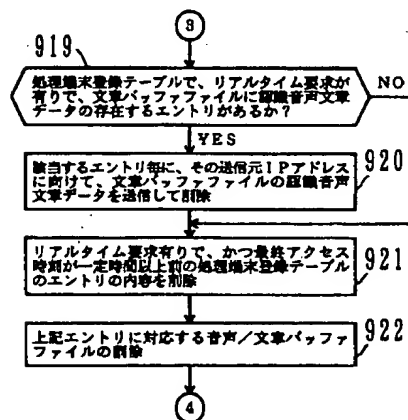
第1ワード	バージョン	ヘッダ長	サービスの種類	IPプログラムの全長
第2ワード	識別番号			フラグ列 フラグ外のビット
第3ワード	生存期間	プロトコル	ヘッダのチェックサム	
第4ワード	送信元IPアドレス			
第5ワード	宛先IPアドレス			
第6ワード	IPオプション			パディング

(b) TCPヘッダ

第1ワード	送信元ポート番号	宛先ポート番号
第2ワード	シーケンス番号	
第3ワード	確認応答番号	
第4ワード	ヘッダ長 予約済 flags	ウィンドウ
第5ワード	チェックサム	緊急ポインタ
第6ワード	オプション	パディング

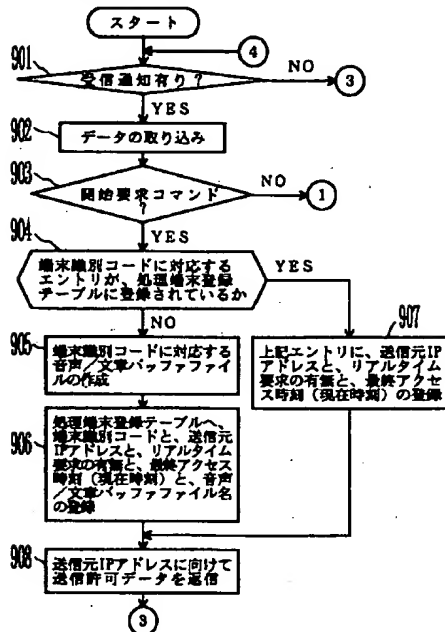
【図12】

移動端末通信制御部の動作フローチャート(その4)



【図9】

移動端末通信制御部の動作フローチャート(その1)



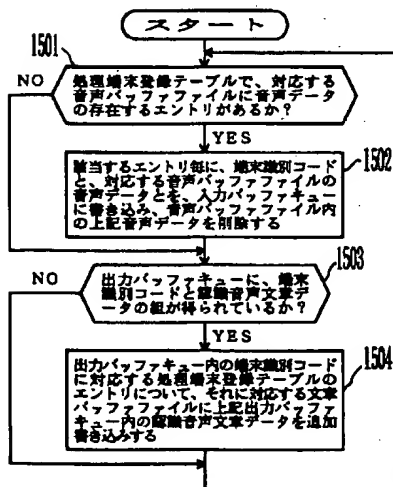
【図13】

処理端末登録テーブルのデータ構成図

文章バッファファイル名			...
音声バッファファイル名			...
音声時刻			...
最終アクセス時刻			...
最終要求有無			...
リアルタイム要求有無			...
送信元IPアドレス			...
端末識別コード			...

【図15】

入出力制御部の動作フローチャート





【図14】

文音声認識部の構成図

